Population and Deforestation in Costa Rica

Luis Rosero-Bixby
University of Costa Rica

Alberto Palloni
University of Wisconsin

This paper addresses a central debate in research and policy on population and environment, namely the extent to which rapid population growth is associated with the massive deforestation currently underway in the tropics. We utilize the experience of Costa Rica during the last forty years to illustrate what the main issues are, discuss the history of deforestation in that country, and present results from conventional regression methods and from the application of spatial analyses. These analyses enable us to estimate the magnitude of the relation between population and deforestation and to identify the factors that are responsible for the linkage between them.

INTRODUCTION

This paper addresses a central debate in research and policy on population and environment, namely the extent to which rapid population growth is associated with the massive deforestation currently underway in the tropics. Although temporal and spatial associations strongly suggest a connection between population growth and deforestation (Preston, 1994), some research indicates that the problem is more complex as it involves

Please address correspondence to Dr. Palloni, Center for Demography and Ecology, University of Wisconsin, 4426 Social Science Bldg., 1180 Observatory Drive, Madison, WI 53706; or Dr. Rosero-Bixby, University of Costa Rica, Apartado 833-2050, San Pedro, Costa Rica.

Population and Environment: A Journal of Interdisciplinary Studies
Volume 20, Number 2, November 1998

non-demographic mechanisms resulting from credit and capital market failures, lack of suitable mediating institutions securing property rights, wretched poverty, uneven land distribution, consumption patterns in developed countries, greedy multinational companies, ignorance and bad management by colonists of frontier land, and so forth (Gillis & Repetto, 1988; Bilsborrow & Ogendo, 1992; Myers, 1984; Palloni, 1994).

This paper is an exploratory analysis of highly disaggregated data from Costa Rica—a tropical country that in the 1960s and 1970s experienced one of the highest rates of deforestation and population growth in the world. It addresses the methodological problem of linking people and population pressure to land cover, a problem that arises from the fact that people usually do not live in the forests that will be cleared. To establish the population-land linkage the paper relies on a multidisciplinary geographic information system (GIS) platform, which was developed for this study with georeferenced data from two population censuses and a series of land cover maps. The key analyses in the paper use multivariate logistic regression to model the net impact of population growth on the 1973–83 probability of deforestation in about 31,000 parcels of 750 meters per side, which were covered with forest at the beginning of the period. Since conventional logistic models fail to account for sources of unmeasured covariates that could cause autocorrelation, we present in an appendix estimates which attenuate the impact of spatially relevant unmeasured covariates.

BACKGROUND AND ANALYTIC FRAMEWORK

Norman Myers, a British ecologist, predicts that in few decades not much tropical forest will remain on Earth, unless there is a marked reduction in population growth and a resolution of the landless-peasant phenomenon (Myers, 1991). This is an extremely worrisome scenario. Although some of the basis for this prediction may be disputed, there is ample evidence that tropical forests are, indeed, disappearing at a very fast pace (FAO, 1990) and common sense suggests a connection between this change and the fast growing population numbers in tropical countries.

Why bother with deforestation at all? Until very recently, clearing the land for cultivation was considered an indication that development and civilization had arrived to the wilderness. Nowadays, however, preserving tropical forests is a well-accepted value. An abundant literature suggests that destroying the forest may be the first link in a chain of environment degradation that includes erosion, climatic changes, loss of biodiversity and genetic endowment, air pollution, decline in watershed functions, and the obvious loss of hardwood, fuelwood, and aesthetic stocks (Myers, 1984; Whitmore, 1990). Most of these consequences are externalities of social processes and activities that markets do not account for, consequently meriting public interventions.

Deforestation is seldom caused by physical phenomena alone. It is mostly a human product. But there is disagreement about the exact role played by population growth and pressure. Some authors who emphasize the demographic dimension underscore as key causes of deforestation the increased need for arable land to absorb excess labor force and keep up with growing demand for food from a larger population, and the increased consumption of fuelwood and timberwood brought about by rapid population growth. Those who minimize the demographic factors portray deforestation as rooted in the political economy, caused mainly by uneven distribution of income, land, and access to credit and capital (Stonish, 1989), rural poverty (Ellen, 1982), international markets that promote and encourage wholesale logging and cattle ranching (Nations & Komer, 1982), market failures due to dysfunctional property rights, bad management and titling policies and, finally, inappropriate and unsuitable technologies (Hecht, 1985). A more nuançed and suggestive view (Moran, 1991) indicates that the impact of human settlement on the frontier is not uniform or homogeneous at all but that, instead, it follows a time-dependent (and possibly space-dependent) trajectory tightly connected to households' composition and life cycle, and just as it can have an initial deleterious effect it may also become if not altogether beneficial at least environmentally neutral.

The slash-and-burn cultivator is often singled out as the most significant agent of deforestation in the tropics (Myers, 1991). He is portrayed as a landless peasant who migrates to the forest to open new agricultural frontiers on public lands. He usually knows little about the forest and its soil, which often results in deployment of inappropriate cultivation techniques and adherence to practices that lead to land degradation. Other agents of deforestation may be farmers who clear their land to cash in on logging or to use the land for cattle pasture or in agriculture for food production, or more importantly, cash crops such as bananas. In both these cases high levels of fertility may be sustained and promoted, thus perpetuating the continued reproduction of an economy geared toward the destruction of the forest.

There are also speculators who clear public lands to claim property rights and sell them later. And, finally, there are the logging companies that exploit forests located on public lands for timber production. However, the effects of the encroachment of these various agents can be vastly different.

In general, replacement of forest by cultivation or pastures totally destroys them, whereas their use for wood production and harvesting (timber or fuel) may partially preserve them (Whitmore, 1990, p. 173). Some researchers suggest that under appropriate mixtures of constraints, opportunities and incentives, preservation of some parts of the rain forest and restoration of other, previously overexploited, terrain may be more easily obtained than we think even in the presence of strong demographic pressures (Moran, 1991).

The framework in Figure 1 sketches some of the causal pathways to deforestation. It postulates two direct connections between population growth and deforestation: (1) relative land shortages in traditional farming areas that result from the combination of growing numbers of peasants, high population density (accumulation of previous population growth), uneven land distribution, and preservation of agriculture technologies favoring extensification over-intensification; and (2) increased demand for timber and fuelwood, which may result in over-exploitation of forests, and increased demand for food with the corresponding need for converting forest lands to agriculture. This paper focuses on the first causal link only: that is, the pressure of growing numbers of cultivators on forest lands. We translate this postulated causal link into a testable hypothesis, namely, that the likelihood of deforestation is higher in forest sites that are in the proximity of populations of cultivators which are larger, growing at faster rates, and more dispossessed. The empirical test we offer is admittedly tentative since it is based on observation of geographic covariations in population and deforestation, not on what is certainly a more appropriate source of evidence, namely, the trajectory over time and space of the relation between patterns of human settlement and patterns of land utilization and degradation.

The second direct pathway to deforestation—increased demand for land products—cannot be properly studied with the data available to us. National and international markets of food and timber blur geographic covariations at the level of aggregation considered in this paper. For example, the increased need for food and timberwood in a city may cause deforestation in the most remote and diverse locations in the country. We do attempt, however, to estimate the association between deforestation and the magnitude and growth of the population that uses fuelwood for cooking even though we are not able to completely resolve delicate issues pertaining to direction of causality.

Population growth is by no means the only direct causal agent of deforestation. Figure 1 postulates four additional factors implicated in tropical forest depletion even in the absence of population pressures. International

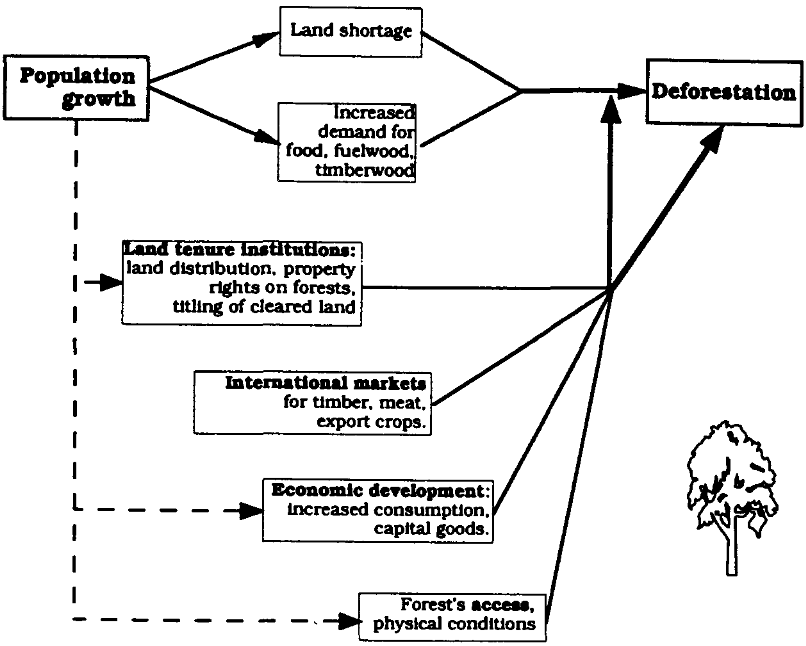


FIGURE 1. Causal diagram of population and deforestation.

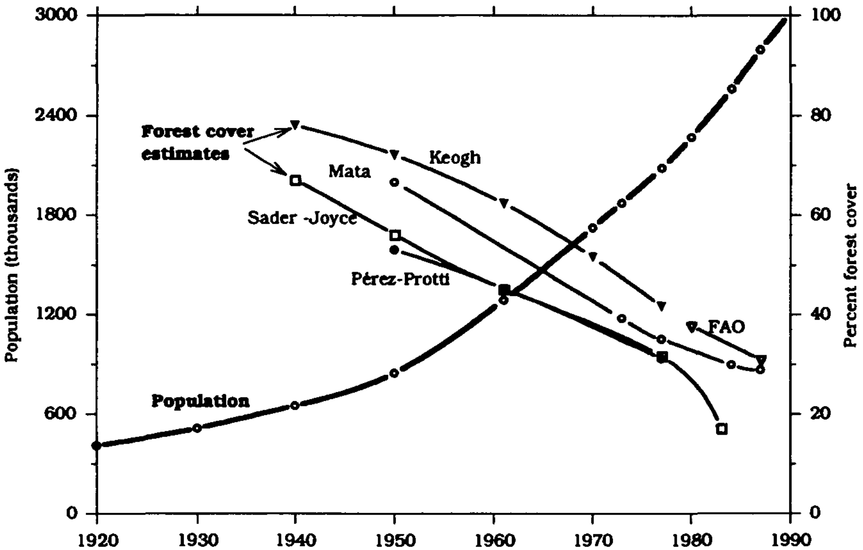
markets and local credit policies favoring banana plantations and cattle ranching (the "hamburger connection") are Costa Rican examples of these processes. The opening of new roads through or near tropical forests and physical conditions of terrain and climate determine the accessibility to forests and, consequently, their probabilities of survival. Increased per capita food and timberwood consumption and improved tools for logging are also potential factors of deforestation brought about by economic development. Property rights on forest covered land and titling policies rewarding forest clearing are seemingly important contributing factors often mentioned in the literature. Although we are not able to assess the independent contribution of these other factors, we consider them as sources of potential unmeasured heterogeneity inducing spatial relations among geographic locations and confounding the relation between population measures and deforestation.

Some of the aforementioned factors may not only have a direct impact but could also exacerbate or attenuate the deforestation consequences of population growth by interacting with demographic effects. For example, the increased demand for labor in manufacturing and the service sector as well as agricultural intensification may absorb landless peasants, thus diverting the effects of population pressure away from the rain forest; highly uneven distribution of arable land may aggravate land shortages with the consequent increase in pressures for population displacement that encourages the search for new frontiers; and titling policies or the construction of new roads can translate population pressure into actual encroachment, settlement and destruction of forest cover. These interactions and synergisms complicate the task of isolating the independent effect of population. Finally, it should be noted that these factors are, to some extent at least, influenced by population growth (dashed arrows in Figure 1). The corresponding causal paths are mediations of the population-deforestation connection: roads are often built because of the population growth in the vicinity of the area they will serve; land fragmentation is the outcome of population pressure exerted within the boundaries of a particular land tenure system; and economic development may be inhibited by rapid population growth. Since these indirect population effects are ignored in the present analysis, the population effects that we estimate are gross rather than net effects.

THE SETTING: COSTA RICA

Costa Rica is an ideal setting for studying the impact of population growth on land cover. This country has rich and accessible data sets on population, land use and other intervening variables for the last three decades. Its relatively small size (about 20,000 square miles) facilitates the manipulation of computer images for the whole country in desktop computers. The country also has one of the greatest diversity of life zones in a small territory in the world. In recent years, the Costa Rican government has been a world leader in the efforts to preserve the environment, which means that the results of this and other studies may serve to shape policies. Most importantly, land use and population size and composition in Costa Rica went through dramatic changes during the study period, an ideal situation for testing hypotheses about the impact of rapid population growth.

Both deforestation and population growth radically changed the Costa Rican landscape in the present century, particularly after World War II. A staggering four-fold increase in the total population, from less than 800,000 to more than 3 million people, occurred in the less than two generations of the post-war era. In the same period, about 50% of Costa Rican territory was cleared of its primary forest cover (Figure 2).



Sources: FAO (internet data base); Keogh 1984; Lutz (Mata) 1993; Pérez - Protti, 1978; Sader - Joyce 1988.

FIGURE 2. Population and forest cover estimates in Costa Rica, 1920–90.

Rapid population growth was the consequence of successful public health programs that dramatically reduced mortality rates. Costa Rica is known for being one of the success stories in the third world, having managed to reach health levels comparable to industrialized countries in spite of its under-developed economy (Halstead et. al., 1985). Life expectancy at birth, for example, was 72.6 years in 1980. Declining mortality rates took the population growth to a peak of almost 4% natural increase per year by 1960: one of the fastest in the world. Although birth rates plummeted after 1960, population momentum kept adult population and the number of households exploding at rates well over 3% per year until the 1980s, when the first cohorts born at lower birth rates started to reach adulthood. Costa Rica, with about 3.5 million inhabitants in a territory of 20 thousand square miles, is nowadays the third most densely populated country in the mainland South and Central American continent (only El Salvador and Guatemala have higher population densities in the continent).

A massive loss of forest cover paralleled the demographic explosion in this country. The details of deforestation trends are, however, blurred by somewhat contradictory estimates (figure 2). By 1940, estimates of primary forest cover of Costa Rica range from 68% (Sader & Joyce, 1988) to 78% (Keogh, 1984). The most recent estimates range from 17% forest cover in 1983 (Sader & Joyce, 1988) to 31% in 1990 (FAO). Part of the discrepancy is simply a matter of definitions. The lowest estimates include in the definition only undisturbed forests; the highest usually include secondary and severely disturbed forests. In any event, all estimates show very high rates of deforestation which, like adult population growth, peaked in the 1970s. Approximately 4% of forest-covered land, or more than 1% of the Costa Rican territory, was cleared each year—one of the highest deforestation rates ever observed. After 1983, however, the massive deforestation process seems to have stopped or even reverted (González, 1993). Bonilla (1985, p. 51–52) describes the deforestation of Costa Rican territory in the following terms:

> Deforestation started slowly in the XIX Century. In 1800, population density was one inhabitant per square kilometer, but with the natural increase of population, forests start to be cleared to convert the land to agriculture. The process started in the central part of the country, where primary forests were replaced by coffee plantations. . . . By the beginning of the XX Century, the agriculture lands in the Central Valley reached a point of saturation and colonization of the rain forest started in a massive way. The resulting flow of settlers eliminated large areas of natural forests. In addition, at that time the cultivation of bananas started, bringing deforestation to large areas in the Atlantic region and, later on, in the fertile lands of the Central and South Pacific regions. . . . When colonists left the Central Valley, they shifted to cattle ranching and extensive agriculture, using large areas of land to support non-dense populations. . . . In this way, cattle ranching impoverished and killed the country. Most of the cleared forest was not even used as timber. . . .

Bonilla is not alone in his claim that population growth is a key factor for deforestation in Costa Rica. The most commonly mentioned causal link between these two processes is the demographic pressure on land combined with public policies favoring settlement in public lands to avoid land reform and to take away population pressure (Hartshorn, 1983; Pérez & Protti, 1978). Some authors also mention indirect causal links, such as ". . . increasing profitability of commercial agriculture, both by the lower cost and greater availability of labor, and the expansion of the domestic market for food and wood" (Harrison, 1991, p. 92).

Other researchers agree with the idea that population growth is not the only cause of deforestation. Among the other factors blamed for Costa Rican deforestation are the boom of banana exports and cattle ranching (mostly driven by international markets), land tenure institutions, government policies, income distribution, relative prices, and wasteful logging technologies (Arcia et. al., 1991; Jiménez, 1991; Kishor et. al., 1993; Lutz et. al., 1993; Sader & Joyce, 1988). In particular, property laws that encouraged land clearing and speculation are at the top of the list of factors contributing to deforestation during the 1960s and 1970s.

In spite of the temporal coincidence of massive deforestation and the population explosion, the empirical evidence of a link between the two processes is surprisingly scant. The causal connection between rapid population growth and deforestation has often been taken as a matter of common sense and relatively little research has been conducted to prove it. A study conducted by Susan Harrison (1991), a biologist, is one of the few studies assessing the deforestation impact of population growth in Costa Rica. Harrison analyses the covariations in population and forest cover for the 65 Costa Rican "*cantones*" existing in 1950. Her analysis applies to the period between the census years 1950, 1973 and 1984. Harrison conducts separate analyses for three regions. Her most relevant results are those for the "frontier region," which comprises about 90% of forested areas in Costa Rica. Unfortunately, this region has only 12 *cantones*, a serious limitation in the "power" of the sample to detect statistically significant associations. Harrison does not find conclusive evidence of a connection between population growth and deforestation. In the frontier region, correlation coefficients are positive but seldom statistically significant. In the other regions, the signs of the coefficient shift erratically and only few estimated effects are statistically significant. Note that these are all results of estimating associations between contemporaneous changes in population size and forest cover and assume the absence of lagged effects. The correlations for levels of population density and forest cover are negative in all three years and regions studied, but most of them are not statistically significant. In light of these results, Harrison asks the most important question, an unanswerable counterfactual: "How much less deforestation would have occurred in the absence of population growth?" She provides an answer consistent with her results showing little evidence pointing to population pressure as the main culprit:

> Relatively few people are required to cause a great deal of deforestation. . . . The economic and other factors could have brought about a great deal of deforestation, even if the popula-

> tion had remained at its 1950 size. . . . Perhaps the destruction of Costa Rica's forest could be said to ultimately have been caused by its status as an open-access resource. . . . Deforestation may have been inevitable as long as this was the case, with population growth only one of many pressures acting to hasten it (Harrison, 1991, p. 91–92).

A more recent study by the World Bank based on the observation of a non-random sample of 52 deforestation sites, also suggests that direct links between population growth and deforestation are weak, at least in the Costa Rica of the 1990s (Lutz et. al., 1993). This comes from the observation that "small holders squatting on public or private land seem to play only a minor role in current land clearing or logging." Or that: "forest clearing to establish a stronger claim to the land no longer appears to be a motive, as it was in the past" (p. iii). Deforestation in the 52 sites surveyed reflects a well organized, highly capital-intensive industry driven by economic considerations of land owners or transnational corporations associated with timber harvesting rather than by local demographic pressures to open new lands for landless migrants. The study concedes, however, that these patterns may not explain the past destruction of forest cover.

In contrast to the conclusions drawn in the World Bank study, an analysis of census and administrative data on settlers and squatters by Cruz (1992) finds that migration of squatters to dense forested areas increased in the 1980s. Moreover, the economic crisis during these years reversed long-term migration trends previously dominated by urban ward flows: frontier ward flows increased and flows away from the capital city toward distant rural areas emerge as an unprecedented phenomenon. As a result, "forests and marginal lands are now increasingly colonized by landless peasants" (p. 3), which, the author suggests, is the major cause of environmental degradation in Costa Rica.

The findings from the studies carried out by Harrison and the World Bank are far from conclusive. Harrison's study is hamstrung by an aggregation problem of some severity since *cantones* in the frontier region are too large, too heterogeneous, and too few to provide a basis for robust inferences. On the other hand, the World Bank study selected a sample of 52 deforestation sites based mostly on information provided by the authorities in the General Directorate of Forestry, which probably biases the results by magnifying large scale and legal logging. In addition, neither of these studies address the key issue of the circumstances that mediate or alter the relation between population and deforestation, including land fragmentation, property rights, titling policies, and alternative employment opportunities.

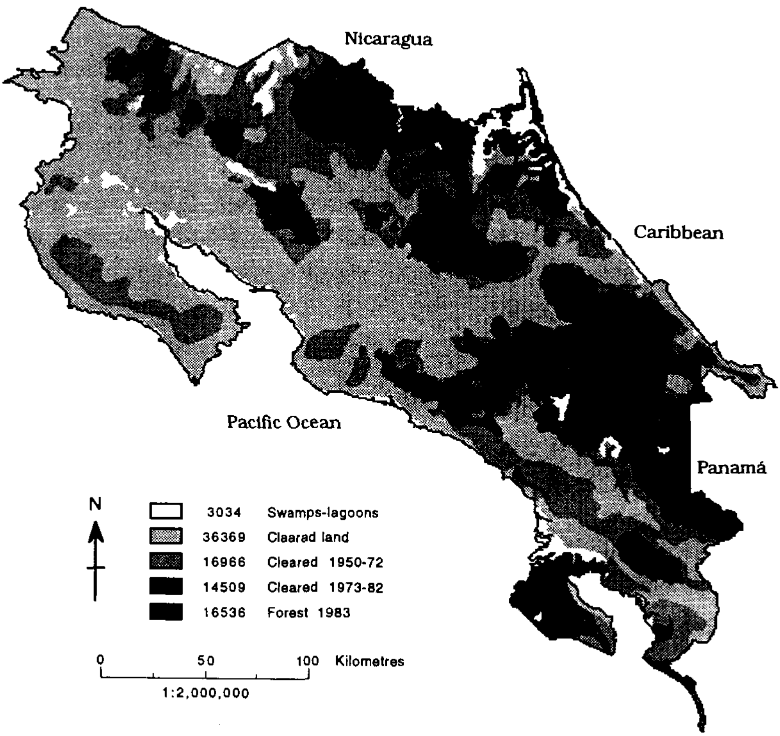
THE DATA: MEASUREMENT ISSUES

We first developed a consistent GIS platform for the whole country with three sets of map layers: (1) land use for a series of years, (2) physical elements including roads and life zones, and (3) population size and characteristics in 1973 and 1984. It was not possible to develop a fourth set of GIS layers on land tenure and production relations because, to preserve the confidentiality of the data, the Census Directorate did not permit access to individual records of agriculture censuses. The GIS platform initially combined raster-based images on land use and physical characteristics with vector-based data on population. The GIS facilitated identification of deforested sites, computation of distance-based population-potential indicators, and visualization of broad patterns. Square parcels of about 750 meters per side were taken as units for statistical analyses. This cell resolution is a compromise between the magnitude of the error in the process of geocoding censuses and digitizing maps, the need for disaggregation, the computing capabilities available, and the resolution of some of the original maps. The Costa Rican territory comprised about 90,000 of these parcels, but statistical analyses were restricted to the about 31,000 parcels covered with forest in 1973 (Map 1). What follows is a description of the three sets of data layers.

Land Use and Deforestation

We use a series of forest cover maps for 1950, 1961, 1977 and 1983 assembled by Sader and Joyce (1988). The aforementioned studies by Harrison (1991) and Keogh (1984), as well as several others, have also used this map series. Low resolution computer images of these maps were downloaded through Internet from the United Nations Environment Program/Global Resource Information Data Base (UNEP/GRID) in Geneva. The original maps were published and developed by the Costa Rican Ministry of Agriculture by interpretation of aerial photos (1950–61) and LANDSAT images (1977 and 1983). The series 1950–1977, published in 1978, is internally consistent. The accuracy of the original maps is, however, unknown. There are probably improvements in accuracy over time. The map for 1977, for example, benefitted from field validations (Sylvander, 1978), which were not possible for earlier maps. The map for 1983, published in the same year, contained more information but also presented some inconsistencies with the earlier series.

Following Sader and Joyce (1988, p. 12), we consider in the study only "primary forests," i.e., relatively undisturbed natural forests with an upper canopy covering more than 80% (90% in 1983) of the surface area.



MAP 1. Deforested land in Costa Rica, 1950–83.

Deforested areas in 1973–83 were identified by comparing 1973 and 1983 map layers (Map 1). Deforestation thus includes conditions ranging from complete removal of forest cover to removal of a few percentage points of the upper canopy. Areas with 80% to 89% forest cover in 1973 that were undisturbed in 1973–83 were misclassified as deforested during this decade because of the classification change in the 1983 map.

We modify the original map layers as follows:

- The 1983 layer was slightly corrected to fit the 1950–77 grid using "rubershed" techniques. The largest corrections, of about one kilometer, were on the Northeastern and Southeastern edges of the map.
- No forest was allowed in areas that according to the 1983 map were swamps, mangroves and lagoons.
- Following Sader and Joyce (1988), no forest regrowth was allowed in this map series, i.e., forest covered areas in later maps had to be also

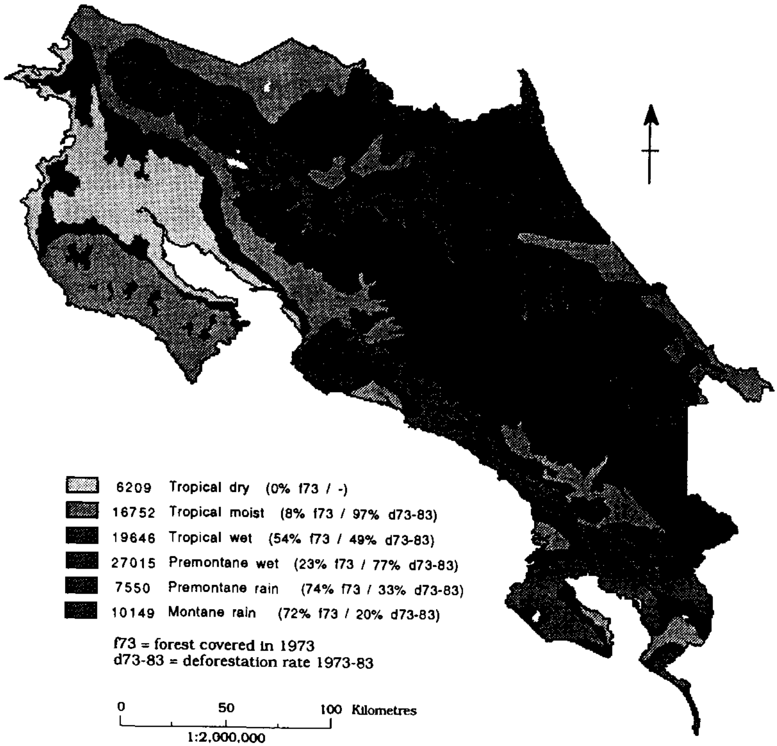
forested in earlier maps (this also assumes that later maps are more accurate).

- The 1973 layer was spatially interpolated by breaking deforested areas in 1961–77 into the 1961–73 and 1973–77 sub periods. In order to split the period, a deforestation trend surface was first estimated with a roving window of 10 kilometer radius on a map showing five possible deforestation periods: before 1950, 1950–61, 1961–77, 1977–83, and 1983 or later. The roving window estimated the likely year of deforestation for its center cell as the simple average of the cells in the window. Among the cells originally in the period 1961–77, the 25% with the latest deforestation years were assumed to be cleared in 1973–77. These areas were added to the 1977 forest cover map to estimate the 1973 map layer.
- Isolated patches of forest and deforestation (smaller than 4 km^2) were excluded using a "clump and sieve" procedure.

We are not completely comfortable with the accuracy of our deforestation estimates. There are uncertainties about the dates of the source material used in the original maps and about the precision of these maps. The change in the classification criteria from 80% forest cover in 1973 to 90% in 1983 is an obvious source of misclassification of some parcels. Interpretation of LANDSAT images has also a margin of inaccuracy (91% accuracy identifying dense forests, according to an USAID study, 1979). The corrections and interpolation described above probably introduced additional errors. Further research should give priority to improving deforestation measurements.

Map Layers of Physical Features

Landscape patterns, including temperature, precipitation, terrain, and accessibility, were brought to the analysis throughout a map of the ecological zones in Costa Rica. The map, following the Holdridge Life Zone System, was originally developed by the Tropical Science Center in Costa Rica (Tosi, 1969) and digitized by Sader and Joyce (1988). The map was downloaded from the UNEP/GRID data base in Geneva along with the land cover images. We combined the original 17 life zones into the six categories shown in Map 2. Premontane and montane cloud forests are usually less accessible and less desirable for agriculture because of high rainfall, rugged terrain and soil fertility limitations (Sader and Joyce, 1988, p. 15). Drier life zones have more favorable climate and soil conditions for agriculture and pasture use, which, consequently, put them at higher risk of deforestation. The consequences of deforestation are also likely to change



MAP 2. Lifezones and forest cover in Costa Rica.

across life zones, since plant biomass correlates with temperature and precipitation conditions (Brown & Lugo, 1984).

Two accessibility surfaces were also included as GIS layers: (1) a surface with the shortest distance of each parcel to a national road according to a 1977 road map (source: Costa Rican Public Works and Transportation Ministry) digitized by Sader and Joyce (1988) and downloaded from the UNEP/GRID data base; (2) a surface with the shortest distance of each map cell to the forest frontier in 1973 (Map 1). Land that is located closer to highways or to forest outskirts is at a higher risk of deforestation. This land is probably also under higher population pressure.

Population Map Layers

We geocoded the 1973 and 1984 population censuses to link them to deforestation and land use data in the GIS platform. The ideal would be to have Earth coordinates for every household in the census. This, of course, is not feasible. We geocoded census tracts ("segmentos censales") instead and represented all households in the tract by a single point in the demographic centroid of this area. Census tracts in Costa Rica contain on the average about 70 households in 1973 and 50 households in 1984. In urban areas the tract usually consists of one or two city blocks. In rural areas it usually is in the range of 5 to 10 square kilometers. There are not important errors in representing all households of a tract with a single point, particularly considering that typical tracts in rural areas contain just one or two clusters of households and a large, empty territory of farm land, making the demographic centroid a good representation of households' location.

There are about 5,000 census tracts in 1973 and 11,000 in 1984. We geocoded them by marking their centroids on census maps and reading the corresponding coordinates. Since Costa Rican census maps do not have Earth coordinates, we linked them to charts of the National Geographic Institute at scales 1:50,000 and 1:10,000 using a landmark in each census map. The accuracy of this geocoding procedure was validated on a probabilistic sample of 40 tracts. The geographic coordinates for this sample were taken in the field with a Global Positioning System (GPS), a device based on satellite signals. The median discrepancy between the two geocoding procedures ranged between 15 and 900 meters, with a median of 60 meters. Considering that GPS-based measurements are not error free, this validation suggested that the error in the great majority of our map-based measurements is less than 200 meters, and that the probability of having errors larger than 500 meters is nil. In this and other computations we projected Earth coordinates to a plane using the North Lambert Conformal projection for Costa Rica (Inter-American Geodetic Survey, 1950).

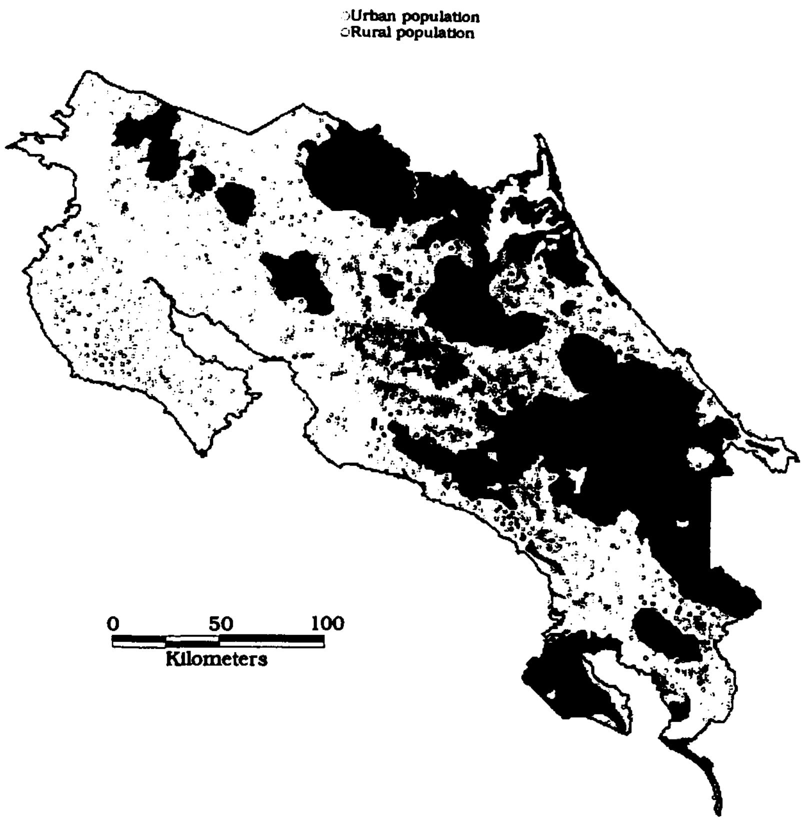
The original maps and the complete data files of the 1973 and 1984 census were made available to this study by the Costa Rican Directorate of Statistics and Censuses. From the original 1973 and 1984 census data files we tabulate at the tract level the following variables for use in the analysis:

- Total population.
- Agriculture population: adult men occupied in agriculture activities.
- Land owners: agriculture population working on their own land.
- Agriculture employees: agriculture population working for a salary.
- Landless peasants: agriculture population who do not own the land and work on their own or as an unpaid family member.
- Total households.
- Households under the poverty line, basic unmet needs criteria (lack of at least two of the following items: running water, toilet, a separate

cooking room, electricity, non bare-dirt floor, radio, and three or less persons per bedroom).

- Households using fuelwood for cooking.
- Net reproduction index: living children per woman aged 40 to 49 years.

This information, combined with the tract's geographic coordinates, conformed a set of map layers in vector-point format in our GIS. These layers, however, are not appropriate for the analysis since only a small fraction of the land contains population information (Map 3). To link land parcels to population tracts we turn to an old friend in demography: the concept of population potential, set forth by Stewart in 1947 (Duncan, 1959, p. 692). The population potential in a land parcel i is given by Σ_j (P_j/D_{ji}), where P_j is the population (total or in a sub-group such as farmers)



MAP 3. Population and deforested land, Costa Rica 1973.

in tract *j*, D_{ji} is the distance between *i* and *j*, and the summation is over all tracts j. We restrict summation to all tracts within a radius of 15 kilometers (10 miles) and compute the corresponding potentials for all of the aforementioned population variables.

The population potential in a land parcel measures the relevant demographic pressure over that parcel. We also compute 1973–84 annual population growth rates for each parcel as the ratio between the population potential change and the population potential average in the 1973–84 period (divided by 11 and expressed as a percentage). Both population potential and population growth rates were included in our GIS as continuous surfaces in a raster-format with cell resolution of 750 meters. In this fashion we had a common unit of analysis for all physical, socioeconomic and demographic data. We explored, and abandoned, three alternative methods for converting population data collected for discrete spatial units into continuous representations, namely: (1) Thiesen/Voronoy polygons (Haggett et. al., 1977); (2) trend surfaces derived with local regression techniques as implemented in the S-plus computer package (Chambers & Hastie, 1992); and (3) an expansion method used to reconstruct settlement geography from georeferenced population points in the 1981 British Census (Bracken & Martin, 1989; Martin & Bracken, 1991). Our choice of the population potential method was based more on practical than on theoretical consideration. It remains to be explored the degree to which our inferences are sensitive to the method we chose.

RESULTS

General Patterns

The deforestation analysis in this paper is restricted to land covered by primary forest in 1973, which represents 36% of Costa Rican territory: about 31,000 parcels each of 750 meters by side (Map 1). Almost one half (47%) of this land appears cleared in the study period 1973–83. This amount of deforestation is extremely high for a period of 10 years. It implies that a clearing rate of about 1.5% of the Costa Rican territory, or 820 km^2, every year.

Estimates of deforested land area for the late 1970s and early 1980s range between 370 km^2/year (Lutz et al., 1993: Table 4.2) and 1,240 km^2/year (Sader & Joyce, 1988) with most estimates hovering in the vicinity of 600 km^2/year (Sylvander, 1978; FAO, 1990; Pérez & Protti, 1978; Hartshorn, 1983). This paper's estimate is thus somewhat high, probably

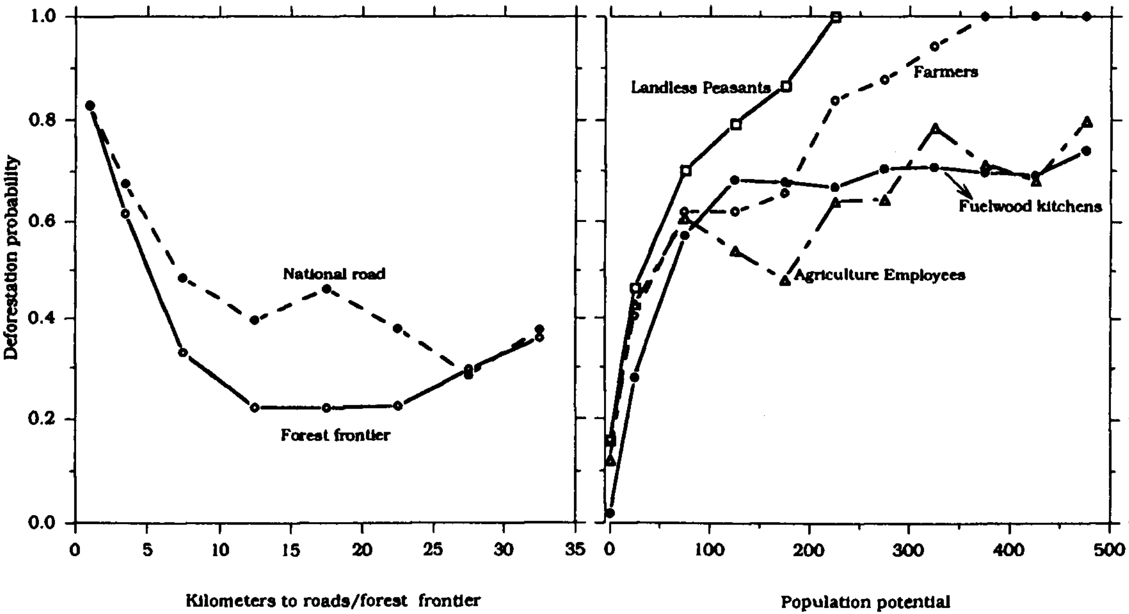


FIGURE 3. Probability of deforestation by distance to roads and forest's frontier, and by population potentials. Costa Rica, 1973–1983.

because of the more demanding definition of primary forest used in the 1983 map (90% coverage) than in earlier maps (80% coverage).

Forest clearing probabilities vary substantially across life zones, from 97% in tropical dry zones to 20% in montane rain (Map 2). The lowest clearing rates correspond to less accessible, less desirable land for agriculture. Deforestation probabilities are also strongly associated with accessibility (Figure 3). More than 80% of the area located near the forests' edges or near (< 2 km) roads was deforested. The risk of deforestation diminishes quickly when one moves a few kilometers away from the forest edge and roads, and levels off after about 15 kilometers. This pattern of diminishing marginal effects at longer distances is modeled in our analyses by transforming distances into their natural logarithms.

Univariate or Crude Effects on Deforestation

Map 3 shows clear evidence that population and forests do not get along at all. The map shows the location of the 1973 population, as represented by census tracts (each dot stands for approximately 400 people), and forest covered lands. Almost no people appear living in forests and no forest exists in populated areas or in their vicinity. The few cases of forested land with human settlements in 1973 were cleared in the following decade.

There is a strong association between measures of population potential and the probability of deforestation: for the lowest value of population potential the probability of deforestation hovers around .16 and then increases monotonically until it is nearly four times as large for the highest population potential. The observed pattern for the total agriculture population potential is reproduced for the subgroup of agriculture employees and for the number of fuelwood-depending households (Figure 3). The gradient of deforestation probabilities is, however, steeper for sub-populations of landowners and, especially, of landless peasants. Within these two subgroups, population potentials large enough may result in 100% forest clearing. As verified with the indicators of distance, the observed association between population potential on deforestation suggests the existence logarithmic effects.

The population growth rates during 1973–83 also show a strong association with the probability of deforestation during the same period (Table 1). Parcels with *negative* annual population growth rates of 3% or larger loss present deforestation probabilities in the 13% to 34% range. Parcels with moderate or null growth present deforestation probabilities of about 50%. In parcels with very high population growth (8% per year or more) rain forest was cleared in about 70% of them. Negative growth rates in the

TABLE 1

Probability of Deforestation by Growth Rates of Selected Population. Costa Rica 1973–83

		Percent Annual Growth Rate			
Population Group	*Total*	*< −3*	*−2 to 2*	*3 to 7*	*8 & +*
Agriculture labor					
Total					
Probability	.49	.25	.50	.66	.76
(N parcels)	(29714)	(9290)	(11343)	(4451)	(4630)
Employees					
Probability	.50	.34	.55	.62	.72
(N parcels)	(28579)	(9287)	(12802)	(3588)	(2902)
Land owners					
Probability	.49	.18	.50	.59	.76
(N parcels)	(29364)	(6066)	(12902)	(5569)	(4827)
Landless peasants					
Probability	.50	.27	.50	.62	.69
(N parcels)	(28991)	(6710)	(11980)	(6248)	(4053)
Households using fuelwood					
Probability	.49	.13	.50	.58	.76
(N parcels)	(29714)	(5539)	(12160)	(7821)	(4194)

number of fuelwood-dependent households coincides with a very low deforestation probability of 13%.

These associations between population growth and deforestation must be interpreted with a great deal of caution since the existence of reverse causality cannot be discarded outright. Newly cleared land may attract settlers in large numbers or may discourage settlers from using fuelwood as the extraction costs mount. In these cases, deforestation either precedes the establishment of human settlements or prevents the occurrence of continued activities that cause forest destruction. However, the emergence of settlements may, in due course, prevent regeneration of forest cover and thus ultimately contribute to the reproduction of conditions that minimize the survival of forests. In this scenario, the relations between population and deforestation flow in both directions but the timing of the corresponding effects is distinct.

Since it is through migration that reverse causation from deforestation to population growth may take place, we should focus on natural population growth in our effort to assess the impact of demographics on deforestation. We use the 1973 index of net reproduction (living children per rural woman aged 40 to 49 years) as a proxy for the natural growth in the number of young adults during the study period. This index, however, presents two analytical drawbacks. It is undefined for land with zero or little population potential. Indeed, in about 10% of the parcels the index was not computable because of the lack of population. A second drawback is that by 1973 past fertility among women aged 40–49 was still uniformly high. In about 85% of parcels the index was six or more children per women (Table 2). The fertility transition initiated in Costa Rica in the 1960s needs a lag of at least 20 years (approximately by 1980s) to make a difference in the natural growth rates of adult population. Table 2 shows that deforestation rates were relatively low in parcels with net reproduction of 3 or 4 children, compared to parcels with 6 or more children. The two extreme groups, however, display deviant patterns: the probability of deforestation is high (.59) in the few parcels with less than 4 children and very low (.16) in those fewer parcels with an extreme high reproduction of 9 or more living children per woman. Perhaps these deviant patterns are due to the fact that these extreme groups are exceptional in a number of ways. For example, area with traditionally high fertility may have relieved past population pressure by contributing to migratory flows toward the city and these flows are reproduced in time and are difficult to replace by alternative behaviors such as the settling of frontier areas.

The connection between rural poverty and deforestation also presents difficulties of interpretation since the potential links act in opposite direc-

TABLE 2

Probability of Deforestation by Net Reproduction and Poverty Levels. Costa Rica 1973–83

*Net reproduction**	*Probability*	*(N parcels)*	*Poverty***	*Probability*	*(N parcels)*
Total	*.51*	*(27862)*	*Total*	*.49*	*(29280)*
< 4 children	.59	(706)	< 20%	.33	(930)
4	.24	(886)	20–39	.45	(4542)
5	.39	(2584)	40–59	.48	(5658)
6	.52	(10423)	60–79	.56	(7435)
7	.56	(9488)	80–99	.59	(7167)
8	.52	(3217)	100%	.24	(3548)
9 +	.16	(558)			

* Net reproduction = Living children per woman aged 40–49 in 1973.
** Poverty = Rural households under poverty (unmet basic needs).

tions. On the one hand, higher deforestation rates may be associated with poverty since land scarcity and population growth are probably much higher among the poor. On the other hand, lower deforestation rates may occur among the poor simply because they cannot afford either the equipment, capital, or even the abundant supply of labor force to undertake clearing projects. Table 2 reflects these contradictory relations. The likelihood of deforestation increases with poverty, from .33 in areas where less than 20% are below a poverty line to about .59 in areas where between 80 and 99% of the population is below the poverty line. In the poorest areas, however, where the entire population is below the poverty line deforestation affects only 24% of parcels.

Multivariate or Net Effects

The evidence examined so far shows a strong association between population and deforestation but it is not altogether certain that the univariate patterns are unaffected by the influence of other factors on both population and the probabilities of deforestation. To address the problem we estimate a multivariate logistic model that includes controls for a number of potential confounders. In addition to population potential measures we include two indicators of accessibility and a set of four dummy indicators of life zones. The dependent variable of the model is the log odds of deforestation in parcels measuring 750 meters on each side. Since we transformed the variables population and accessibility (distance) into their natural logarithms, the logistic regression coefficients estimated for these log variables measure elasticities, i.e., the expected proportionate change in the odds of deforestation given a one point proportionate change in the explanatory variable. Table 3 shows the estimated coefficients for a simple model specification that includes just one population variable: the number of potential cultivators. Preliminary analyses suggested a complex pattern of statistical interactions in which population effects vary by life zone, accessibility and density. To account for some of these interactions, we stratified the sample in two strata according to population potential densities: one stratum ("low density") consists of all parcels with fewer than 100 cultivators and the other consists of all the parcels with densities more than 100 cultivators ("high density"). To ensure simplicity we ignore other interaction effects. Consequently, the model we estimate leads to an averaging of the population effects across life zones and accessibility levels.

The elasticity of population-deforestation in low density areas is substantial: a one-percent increase in the number of potential cultivators results in 0.37% higher odds of deforestation. This translates into effects on

TABLE 3

Logistic Regression Coefficients on the Probability of Deforestation. Costa Rica 1973–83

Explanatory Variables	*All Parcels*		*Low Pop. Density*		*High Pop. Density*	
	Coef.	*(z)*	*Coef.*	*(z)*	*Coef.*	*(z)*
1973 agriculture population (log)	0.291	(−20.3)	0.371	(−18.0)	−0.023	(−0.5)
Accessibility:						
Km to forest frontier (log)	−1.014	(−43.1)	−0.448	(−14.6)	−1.689	(−42.8)
Km to a highway (log)	0.152	(−5.8)	0.378	(−9.1)	0.069	(−1.9)
Life zones						
Tropical wet	0.000	Refer.	0.000	Refer.	0.000	Refer.
Tropical moist	3.352	(−20.0)	2.980	(−17.4)	5.206	(−5.2)
Premontane wet	1.479	(−36.4)	1.219	(−24.4)	1.373	(−20.5)
Premontane rain	−0.980	(−24.6)	−1.093	(−14.4)	−0.964	(−18.7)
Montane rain	−1.205	(−29.9)	−2.053	(−27.2)	−0.902	(−15.9)
Constant	0.106	(−0.9)	−1.863	(−11.1)	2.880	(−10.3)
N parcels	31,045		16,271		14,774	
Pseudo R2	0.306		0.298		0.270	
1973–83 agriculture population growth (percent per year) entered in the model above	0.091	(−43.2)	0.094	(−38.1)	0.062	(−11.9)

The coefficient of the "log" variables estimates the elasticity on the deforestation odds.

the probability of deforestation that amount to increases of 5, 2 and 1% depending on whether the proportion of deforested areas is .20, .50 or .80 respectively. There is no significant population effect in high-density areas, although, judging by the magnitude of the regression constant, these areas have substantially higher deforestation rates than low-density areas to begin with.

The lower panel in Table 3 displays the regression coefficients for the 1973–83 rate of population growth when entered into the model after other variables have been controlled for. The growth rate appears to have significant effects in both low- and high-density parcels. An increase of one percentage-point in the annual growth rate is associated with a 10% increase (exp(0.091)-1) in the odds of deforestation. As mentioned before, however, this effect may be contaminated by simultaneity biases that we are not accounting for.

It is possible that the effects of population potential are not the same across social groups, particularly if these belong to different locations in the process of production. If so, the models in Table 3 erroneously constrain the effects to be the same across groups. Table 4 shows the estimates obtained when each group's population potential is allowed to have its own effects. We include estimates for three sub-groups of cultivators (landowners, employees, and landless peasants) as well as for the number of fuelwood-dependent households. In addition, we control for ecological area and for net population growth potential, levels of poverty, ecological area and measures of accessibility. The results reveal three important patterns: (1) Neither the numbers of landowners, agriculture employees, nor that of fuelwood-dependent households significantly affect the odds of deforestation. Only the effects associated with the number of landless peasants reveal statistically significant elasticities of 1.0 and 0.2 for low and high density areas, respectively. (2) Poverty shows a significant effect but it does so only in the areas with high demographic density: a one percent increase in the proportion of the population below the poverty line increases the odds of deforestation by about .3%. (3) There is a perverse, negative effect of the net reproduction index in both low and high-density areas. Overall, an extra child per woman reduces the odds of deforestation by 7%. This finding reproduces the odd shape of the univariate effects uncovered in Table 2 and, as suggested there, could be explained by resorting to an association between levels of fertility and the history of migratory processes. However, as we show in the Appendix, the estimated effects are likely to be affected by the influence of unmeasured characteristics.

The lower panel of Table 4 displays the regression coefficients for the population growth rates. They are all statistically significant, but two of

TABLE 4

Logistic Regression Coefficients on the Probability of Deforestation for Selected Population Indicators. Costa Rica 1973–83

Explanatory Variables	*All Parcels*		*Low Pop. Density*		*High Pop. Density*	
	Coef.	*(z)*	*Coef.*	*(z)*	*Coef.*	*(z)*
1973 Population:						
Agriculture employees (log)	0.026	(0.8)	−0.013	(−0.3)	0.076	(1.1)
Land owners (log)	−0.179	(−3.1)	0.052	(0.4)	−0.118	(−1.7)
Landless peasants (log)	0.592	(15.7)	1.009	(16.6)	0.217	(3.5)
Fuelwood kitchens (log)	0.186	(2.3)	−0.193	(−1.2)	0.013	(0.1)
Net reproduction (children)	−0.064	(−4.3)	−0.038	(−2.1)	−0.099	(−3.0)
Poverty (percent)	0.013	(10.9)	0.001	(0.5)	0.026	(14.9)
Accessibility:						
Km to forest frontier (log)	−0.938	(−34.5)	−0.564	(−14.2)	−1.455	(−33.3)
Km to a highway (log)	0.035	(1.2)	0.645	(13.5)	−0.262	(−6.3)
Life zones						
Tropical wet	0.000	Refer.	1.000	Refer.	1.000	Refer.
Tropical moist	3.145	(18.7)	2.608	(14.9)	5.096	(5.1)

TABLE 4 (*Continued*)

Explanatory Variables	All Parcels		Low Pop. Density		High Pop. Density	
	Coef.	(z)	Coef.	(z)	Coef.	(z)
Premontane wet	1.442	(33.7)	1.177	(21.2)	1.248	(17.9)
Premontane rain	−1.099	(−23.9)	−1.125	(−13.5)	−0.836	(−13.5)
Montane rain	−1.099	(−22.1)	−1.961	(−23.5)	−0.526	(−6.9)
Constant	−0.672	(−3.5)	−2.393	(−8.3)	1.686	(4.5)
N parcels	27,862		13,088		14,774	
Pseudo R2	0.301		0.297		0.293	
1973–83 population growth (percent per year) entered in the model above:						
Agriculture employees	0.032	(4.6)	0.079	(9.4)	−0.053	(−3.5)
Land owners	0.178	(15.4)	0.282	(16.9)	0.052	(2.8)
Landless peasants	0.020	(3.6)	0.019	(2.7)	0.021	(2.2)
Fuelwood kitchens	−0.087	(−6.3)	−0.241	(−12.8)	0.150	(6.3)

The coefficient of the "log" variables estimates the elasticity on the deforestation odds.

them are improperly signed: those for agricultural employees in high density areas and for fuelwood kitchens in low density areas. It is possible that these patterns are the result of reverse causation that we are not properly accounting for. Thus, the forest-preserving effect of faster growing numbers of agriculture employees may occur if expanding employment in agriculture "factories" removes the demographic pressure on forest lands. If this is so, it should be the case that faster growth of agriculture factories must take place in areas with high population density since it is there where the effects of the rate of growth of agricultural employees is negative. The negative sign of fuelwood-depending households is probably due to the fact that costs of fuelwood extraction grow disproportionately as the forest is cleared and the population initially depending on it shifts to alternative sources of energy.[1]

DISCUSSION

In this document we utilize a geographic information system (GIS) with data on land use, demographics, and physical features to explore the connection between population and deforestation in Costa Rica during 1973–1983, a period during which the country experienced both explosive population growth and massive rain forest clearing. The analysis focuses on about 31,000 land parcels, of 750 meters to the side, covered with primary forest in 1973. We estimate that 46% of this area lost its forest cover in the ensuing decade, an extremely high deforestation rate.

Maps of population and land cover show an obvious and strong pattern: people and rain forest seldom coexist in the same area. Given that almost no people were present in our parcels of forested land, we compute population potentials to measure demographic pressure on each parcel. We detect a strong univariate association between the measures of population potential and probabilities of deforestation. Parcels with 100 or more potential cultivators are four times more likely to be deforested than parcels with less than one potential cultivator.

Most of these effects persist in low population density lands in a sim-

[1]Although these results are plausible, they must be taken with some caution since the regression estimates rely on the implausible assumption that observations in this data set are independent from each other. Deforestation and other variables are obviously correlated among neighboring parcels; i.e., they are spatially autocorrelated. In the appendix we illustrate the effect of introducing alternative corrections for spatial autocorrelation in the logistic model. Some of the results, particularly those suggesting a population-deforestation link, change dramatically with these corrections, confirming our fears that spatial autocorrelation effects may be distorting at least part of the analysis.

ple multivariate framework, after controlling for accessibility and ecological zones. A one-percent increase in the number of potential cultivators increases the odds of deforestation by about 0.37%. In areas of the rain forest that survive high demographic densities, variations in the number of potential cultivators do not affect the odds of deforestation.

Land tenure and relations of production are important for the population-deforestation connection. The demographic pressure of landowners and farms' wage-workers is not a significant factor for land clearing in this data set. In contrast, the pressure of landless peasants is a significant factor, with a sharper effect in low population density areas.

The data do not show significant deforestation pressure originating in the number of fuelwood dependent households, nor do we find a connection between reduced net reproduction and deforestation rates. The simple model shows a negative relation between net reproduction levels and probabilities of deforestation. The data also showed strong and pervasive statistical associations between the rates of population growth and deforestation. But the interpretation of these effects is difficult without resolving the underlying simultaneity problem identified before.

Although the results from the simple multivariate framework are plausible, some of them change in models that attempt to remove spatial autocorrelation. Indeed, using alternative procedures with two different definitions of contiguity we succeed in showing that only the effects of accessibility and ecological areas remain as they were in the simple models and that the effects of all other variables are considerably reduced. In particular, it is no longer possible to attribute deforestation potential to landless peasants or agricultural employees.

The study period, 1973–83, is too early to show the effect of the fertility transition that started in the late 1960s in rural areas of Costa Rica. The population effects on deforestation, if any, that we document in this paper have been accumulated during several decades of population growth. Built-in population momentum makes birth control a poor option for preserving the rain forest now and in the next few decades. From a conservationist point of view, far more important than establishing a connection between population growth and deforestation is understanding how this connection works; in particular, one needs to identify the factors that exacerbate or attenuate it. Conservation policies could act on these intervening factors to meet the challenge of population growth brought about by the demographic momentum of previous growth.

The identification of areas at highest risk of deforestation because of the mixture of built-in demographic pressures, ecology and accessibility, is also important for policy interventions. Some of the estimates obtained in this paper, as well as the GIS assembled during the course of this research could be useful for risk assessment purposes and to convey information to policy makers.

ACKNOWLEDGMENTS

This paper was presented at the Annual Meeting of the Population Association of America, New Orleans, 1996. We acknowledge the support of NICHD Center grant HD05876 to the Center for Demography and Ecology at the University of Wisconsin-Madison.

REFERENCES

Arcia, G., L. Merino, Mata, A., & O'Hanlon, B. (1991). *Modelo interactivo de población y medio ambiente en Costa Rica 1990.* San José, Costa Rica: Asociación Demográfica Costarricense.

Bilsborrow, R., & Ogendo, H. (1992). Population driven changes in land use in developing countries. *Ambio* 21(1), 37–45.

Bonilla, A. (1985). *Situación ambiental de Costa Rica.* San José, Costa Rica: Ministerio de Cultura, Juventud y Deporte.

Bracken, I., & Martin, I. (1989). The generation of spatial population distributions from census centroid data. *Environment and Planning* A, 21, 537–543.

Brown, S., & Lugo, A. E. (1984). Biomass of tropical forest: A new estimate based on forest volumes. *Science* 223, 1290–1293.

Chambers, J. M., & Hastie, T. J. (Eds.). (1992). *Statistical models. S.* Pacific Grove, California: Wadsworth & Brooks.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models and applications.* London: Pion.

Cressie, N. A. C. (1993). *Statistics for spatial data.* New York: Wiley.

Cruz, M. C. J. (1992). Population pressure, economic stagnation, and deforestation in Costa Rica and The Phillippines. In IUSSP Committee on Population and Environment and ABEP, *Population and Deforestation in the Humid Tropics.* Campinas, Brazil. Belgium: IUSSP-UIESP.

Duncan, O. D. (1959). Human ecology and population studies. In P. M. Hauser and O. D. Duncan (Eds.). *The Study of Population,* pp. 678–726. Chicago: The University of Chicago Press.

Ellen, R. (1982). *Environment, subsistence and system.* Cambridge: Cambridge University Press.

FAO (1990). Forest resources assessment 1990: Tropical countries Rome: Food and Agriculture Organization of the United Nations.

Gillis, M., & Repetto, R. (1988). *Public policies and the misuse of forest resources.* New York: Cambridge University Press/World Resources Institute.

González, L. F. (1993). La cobertura forestal en Costa Rica." *Nuestro Agro,* July 21.

Haggett, P., A. Cliff, & Frey, A. (1977). *Locational analysis in human geography.* New York: John Wiley.

Halstead, S. B., Walsh, J. A. & Warren, K. S. (Eds.). (1985). *Good health at low cost: Proceedings of a conference held at the Bellagio Conference Center,* Bellagio, Italy. New York: The Rockefeller Foundation.

Harrison, S. (1991). Population growth, land use and deforestation in Costa Rica 1950–1984. *Interciencia*, 16(2), 83–93.
Hartshorn, G. (1983). *Costa Rica perfil ambiental*. San José, Costa Rica: Trejos.
Hecht, S. (1985). Environment, development and politics: Capital accumulation and the livestock sector in Eastern Amazonia. *World Development*, 13(6), 663–684.
Inter-American Geodetic Survey (1950). *Proyección lambert para Costa Rica*. Washington D.C: Army Map Service.
Jiménez, W. (1991). Situación actual del recurso forestal en Costa Rica y perspectivas de la ordenación natural de sus bosques. In J. Gracia Bondía (Ed.). *El deterioro ambiental en Costa Rica: Balance y perspectivas*, pp. 25–38. San José, Costa Rica: Editorial de la Universidad de Costa Rica.
Keogh, R. M. (1984). Changes in the forest cover of Costa Rica through history. *Turrialba*, 34(3), 325–331.
Kishor, N. M., & Constantino, L. F. (1993). Forest management and competing land uses: An economic analysis for Costa Rica. (Laten Disemination Note). Washington, DC: The World Bank.
Lutz, E., M. Vedova, Martínez, H. San Román, L. Vásquez, R. Alvarado, A. Merino, L. Celis, R. & Huising, J. (1993). Interdisciplinary fact-finding on current deforestation in Costa Rica. (Environment Working Paper). Washington, DC: The World Bank.
Martin, D., & Bracken, I. (1991). Techniques for modelling population-related raster databases. *Environment and Planning*, A, 23, 1069–1075.
Moran, E. F. (1991). Deforestation and land-use in Brazilean Amazon. *Human Ecology*, 21(1), 1–21.
Myers, N. 1984. *The primary source: Tropical forests and our future*. New York: W.W. Norton.
Myers, N. 1991. The world's forests and human populations: The environmental interconnections." In K. Davis & M. S. Berstam (Eds.). *Resources, environment and population: Present knowledge, future options*, pp. 237–251. (Supplement; *Population and Development Review* 16). New York: The Population Council/Oxford University Press.
Nations, J. D., & Komer, D. I. (1982). *Rainforests, cattle and the hamburger society*. Austin, Texas: Center for Human Ecology.
Palloni, A. (1994). The relation between population and deforestation: Methods for drawing causal inferences from macro and micro studies. In L. Arizpe, M. P. Stone, & D. C. Major (Ed.). *Population and environment: Rethinking the debate*, pp. 125–165. San Francisco: Westview Press.
Pérez, S., & Protti, F. (1978). *Comportamiento del sector forest a l durante el período 1950–1977*. San José, Costa Rica.: Oficina de Planificación Sectorial Agropecuaria.
Preston, S. M. (1994). Population and environment: The scientific evidence. In F. Graham-Smith (Ed.). *Population—The complex reality*, pp. 85–92. Golden, Colorado: North American Press.
Sader, S. A., & Joyce, A. T. (1988). Deforestation rates and trends in Costa Rica, 1940 to 1983. *Biotropica*, 20(1), 11–19.
Stonish, S. C. (1989). Processes and environmental destruction: A Central American Case Study. *Population and Development Review*, 15(2), 269–296.
Sylvander, R. B. (1978). Los bosques del país y su distribución por provincias. (PNUD/FAO-COS/72/013). (Trabajo). San José, Costa Rica: Dirección General Forestal.
Tosi, J. A. (1969). *Mapa ecológico de Costa Rica*. San José, Costa Rica: Tropical Science Center.
USAID (1979). Diseño de un inventario y sistema de información de recursos naturales para Costa Rica. (Contrato AID/1a-C-1253). Washington, DC: USAID, División de Operaciones Regionales América Latin.
Whitmore, T. C. (1990). *An introduction to tropical rain forests*. New York: Oxford.

APPENDIX: CORRECTIONS FOR SPATIAL AUTOCORRELATION

The results discussed confirm, albeit superficially, some hypotheses about mechanisms linking population and deforestation. However, all of them are based on a model that relies on a questionable assumption, namely, that the deforestation outcomes across geographic locations are conditionally independent. There are two mechanisms that could produce a relation between outcomes in contiguous parcels. First, it is very unlikely that the independent variables we control for in the model capture all or even most of those which affect deforestation probabilities. Indeed, at the outset we implicated the existence of social, economic and political conditions that we are not able to measure in this study. If, as seems plausible, these unmeasured characteristics are correlated across contiguous parcels, the simple logistic model can only account for their effects to the extent that they are partially correlated with variables that we do include in the model. Under these conditions, it is very likely that a residual correlation between contiguous spatial units will remain.

There is a second mechanism that will lead to conditionally dependent observations and this can be best construed as the consequence of a diffusion process. To the extent that clearing of forests requires learning, accumulated experience, and adequate assessment and management of risks, it is likely that the occurrence (or non-occurrence) of clearing in one place will lead to the transmission of acquired know-how from that place to neighboring ones, thus increasing (decreasing) the probabilities of deforestation in contiguous areas. Insofar as the variables we include in the models capture the diffusion process only incompletely, our observations will be autocorrelated.

Either of these two processes will lead to a high degree of clustering of outcomes. The technical estimation problem that this generates is that the matrix of variance-covariance of the error term in a generalized linear model can no longer be written down in the conventional tidy diagonal form. In the same way, the existence of autocorrelation violates the assumption of independence of observations invoked to justify the maximum likelihood formulation for the estimation of the logistic model. The consequences of this will be the same: we will either produce inconsistent estimates of the coefficients or inconsistent estimates of the standard errors or both simultaneously.

If the dependent variable were continuous and defined on the time domain, the solution would be to use well known autocorrelation models. We will use analogous procedures specially formulated for spatial analysis but merely as a diagnostic tool since they are solely applicable to continuous, not discrete outcomes as ours is.

Another strategy is to formulate the problem as one of clustering of outcomes within well-defined 'families,' 'geographic neighborhoods,' 'contiguous places' or 'clusters' and remove the correlation within clusters with procedures designed to suppress the influence of unmeasured characteristics that inflate intra-cluster correlation. Below we use two different variants of these procedures that differ in the assumptions about how benign the process of intra-cluster correlation is.

We address three different issues: a) what defines a cluster?, b) is clustering of outcomes present in the data?, and c) how can the effects of such clustering be detected and, if possible, removed?

Defining Spatial Contiguity

Models for autocorrelation in the time domain require us to define the number of time units beyond which relations between error terms cease to be important. In the spatial analogue we need to define the boundaries of contiguity or the set of units considered to be neighbors or contiguous to any index case. Since we have no straightforward theoretical directive for doing this, we apply two alternative and somewhat extreme definitions and test the extent to which the estimates are robust to changes. According to the first definition (Criterion I) any parcel (index case) i belongs to a neighborhood of contiguous cases defined by at most 4 parcels in each of the four possible directions in a two dimensional grid, provided that the distance between the center of the index case and any contiguous parcels does not exceed 12 km. The second definition (Criterion II) is less restrictive and places the boundaries of contiguity to 20 parcels in each of four directions and increases the distance requirement to maximum of 60 km.

Naturally, other definitions of neighborhood are possible. However, our intent here is not to obtain exact estimates but rather to assess the degree to which results obtained with conventional models are sensitive to violations of assumptions.

Assessing the Magnitude of Spatial Autocorrelation

The first exploratory tool to detect clustering of results is the calculation of a simple joint count statistic, F, defined as follows:

$$F = .5^*\Sigma_{i,j}\, w_{ij}{}^*(y_i - y_j)2$$

where y_i is a dummy indicator for the outcome of interest (deforestation or not) in the index parcel, y_j is the dummy indicator for deforestation in parcel j, and w_{ij} is a dummy variable attaining the value 1 whenever the pair (i,j) are contiguous neighbors. In the absence of clustering we would expect F to have a value equal to:

$$E(F) = .5^*p^*(1-p)^*\Sigma_{i,j}\, w_{ij}$$

where p is the observed proportion of deforested cases in the total sample. Given our definitions of contiguity and the proportion of deforested units in the sample, the maximum value that F can attain is 2 under criterion I and 10 under criterion II. After calculating the estimates of the standard deviation for F we compute a z-score statistic which, under large sample properties, is normally distributed. The values of the test statistic are .98 when criterion I is used and 5.15 when criterion II is used. In both cases we reject the null hypotheses of no spatial autocorrelation. This confirms results (not shown here) we obtain when using well-known measures suitable to detect intra-cluster correlation on continuous outcomes. Thus, regardless of whether we use discrete or continuous statistics and irrespective of the criterion used to defined cluster, the conclusion is always the same: there is a fair amount of intra-cluster correlation.

The fact that the joint-count statistic (or its analogue for continuous variables) confirms the existence of spatial autocorrelation does not by itself indicate that the estimates of the conventional logistic model are inconsistent. In fact, what is important is whether there is any residual spatial autocorrelation *after* controlling for relevant covariates. Verification of this is straightforward when modelling a continuous variable defined in the time domain. Indeed, all we need to do is calculate a statistic on the residuals associated with the main model and use it to check the persistence of (temporal) autocorrelation. In the present case such strategy could be employed but only in a cumbersome and somewhat arbitrary way since the dependent variable is a discrete, not a continuous, outcome. To circumvent this difficulty we employ a number of alternative techniques, none of which is satisfactory by itself but which taken as a set may shed light on the robustness of conclusions from the conventional logistic model.

Alternative Strategies

The first procedure is designed in analogy to the unrestricted autocorrelation model (Cressie, 1993; Cliff & Ord, 1981) and requires the estimation of the following equation:

$$y - rWy = Xb - WXg$$

where y is the vector of log odds of deforestation, X is a matrix of covariates, b is a vector of effects and $g = rb$. W is a matrix of weights where the values applying to any pair of units (row-column combination) (i,j) are 0 for all parcels j that are not contiguous to the index parcel i and $1/d_{ij}$ for those parcels j that are contiguous to parcel i and are located at a distance d_{ij} from i. It should be the case that d_{ij} is less than the upper limits determined by criteria I or II. Thus W depends on the definition of contiguity and ensures that only contiguous cases contribute non-zero values. To define W we need to identify rules to generate a cluster for each unit in the sample. The cluster to which case i belongs is defined as the set of k closest cases located in each of four directions (North, South, East and West). Thus, each case in the sample—except those on the edges of the area we study—will belong to clusters containing about k^2 cases. In accordance with our previous definition of contiguity we assign k the values 4 and 20.

The estimates of the resulting two alternative models are displayed in Table A1. The first column of the table shows the estimates when $k=4$ and the second column shows estimates corresponding to $k=20$. The most important features of this table are the following:

a) Irrespective of the value of k the effects of number of landowners becomes positive and statistically significant whereas the estimated effect for landless peasants becomes negative and statistically significant. This is in partial agreement with the results of Table 4 obtained for low and high density populations.

b) The estimated effects of poverty become stronger and retain statistical significance: a one percent increase in the proportion of the population below the poverty line increases the odds of deforestation by about 2 to 3 percent (as opposed to one percent in the original model).

c) The perverse effect of the net reproduction rate disappears and becomes statistically insignificant.

d) All other estimated effects remain unchanged.

Thus, this first adjusted model procedure suggests that the estimated effects of the relative size of the two social classes are very sensitive to intracluster correlation,

TABLE A1

Model with Adjustments for Spatial Autocorrelation Cliff-Orcutt Type of Correction (t values in parentheses)

Variables	*Adjusted Model(k=4)*	*Adjusted Model(k=20)*
1973 population		
Agricultural Employees	−.011 (.059)	.179 (1.88
Land Owners	1.633 (5.74)	1.616 (16.08)
Landless Peasants	−.125 (.423)	−.415 (3.28)
Fuelwood Kitchens	.273 (1.77)	−.109 (1.68)
Net Reproduction Rate	.055 (.64)	.051 (1.26)
Poverty	.018 (2.21)	.028 (9.74)
Km to forest	−.601 (2.72)	−.831 (11.10)
Km to highway	.220 (.958)	.121 (1.79)
Tropical wet	1.032 (1.57)	1.55 (5.82)
Tropical moist		
Premontane wet	.198 (.303)	−.020 (1.84)
Premontane rain	.442 (.400)	−.06 (.41)
Montane rain	.139 (.490)	.19 (1.71)
Pseudo R^2	.914 .75	
Log Likelihood	−1605	−3777

that the effects of poverty are stronger than previously thought and that the negative effect of net reproduction rates is probably an artifact.

The second procedure is analogous to the so-called Markovian approach employed in the analysis of mortality and morbidity to eliminate intra-family clustering effects. The solution consists of creating a variable Z_i for index case (parcel) i which summarizes the outcomes among contiguous cases. Thus we define Z_i as follows:

$$Z_i = \Sigma_{j \epsilon c}\, y_j / k$$

where y_j is the outcome in the contiguous case $j \epsilon c$ is the set of contiguous cases for i (the cluster) and k is the number of parcels in cluster c. Thus, the value of Z_i will increase with deforestation in nearby parcels and will decrease with lack of deforestation. As before, the selection of the set of j's that pertain to the neighborhood depends on the criteria we apply (I or II).

This is admittedly an arbitrary model which could be replaced by others, equally plausible and justifiable. For example, the model assumes that the outcome in any parcel is equally influential to that verified in any other parcel contained within the same cluster. That is, we assign no special weight to the distance separating any parcel j from the index parcel i. This is tantamount to saying that if clustering is due to unmeasured characteristics, consistent estimates of the effects of covariates will be obtained after we control for the outcomes observed within each

TABLE A2

Model with Adjustments for Spatial Autocorrelation Markovian Type of Correction (t values in parentheses)

Variables	*Adjusted Model(k=4)*	*Adjusted Model(k=20)*
1973 population		
Agricultural Employees	−.118 (1.56)	−.308 (5.84)
Land Owners	.681 (6.00)	.478 (5.57)
Landless Peasants	−.235 (2.72)	−.240 (4.02)
Fuelwood Kitchens	.346 (1.93)	−.081 (.064)
Net Reproduction Rate	−.064 (1.87)	−.051 (2.09)
Poverty	.012 (.72)	.002 (.90)
Km to forest	−.617 (10.78)	−.860 (20.38)
Km to highway	.043 (.700)	−.014 (.31)
Tropical wet	4.060 (17.40)	3.930 (18.19)
Tropical moist		
Premontane wet	−.11 (1.03)	−.240 (3.26)
Premontane rain	−.29 (2.43)	−.517 (6.10)
Montane rain	.460 (4.55)	.634 (8.95)
Pseudo R^2	.847	.711
Log Likelihood	−2951	−5581

cluster, irrespective of the relative location of the contiguous parcels. One could argue that this is inaccurate since the 'closer' a parcel j is to the index parcels, the larger the magnitude of the correlation between (i,j) outcomes due to unmeasured characteristics ought to be and, therefore, the more influential should j's outcome be on the control for the unmeasured influences on i's outcome. Acceptance of this argument calls for a value of Z_i incorporating weights that are inversely proportional to distance.

The results of this model are displayed in the two columns shown in Table A2 which correspond to the two alternative definitions of contiguity. The results lead to modifications that are very similar to those verified in the previous model: the effects of landowners and landless peasants are reversed whereas those associated with demand for wood and the net reproduction rate remain unaltered or at worst lose statistical significance depending of which criterion we use. In addition, the effects of poverty remain the same but lose statistical significance. All other effects remain the same.

This alternative model is a more strict representation of the clustering in the data and represents a more demanding test to the robustness of conventional results than the previous model. It therefore should not be surprising that we obtain more drastic changes when implementing it.

The final correction procedure targets the estimates of the standard errors rather than the estimated effects. Insofar as the sample of parcels is a clustered

TABLE A3

Model with Adjustments Sampling Clustering Huber Type of Correction (t values in parentheses)

Variables	*Adjusted Model(k=4)*	*Adjusted Model(k=20)*
1973 population		
Agricultural Employees	.025 (.772)	.025 (.09)
Land Owners	−.179 (3.13)	−.179 (.53)
Landless Peasants	.592 (15.79)	.592 (1.82)
Fuelwood Kitchens	.186 (2.29)	.186 (.336)
Net Reproduction Rate	−.064 (4.29)	−.064 (.676)
Poverty	.013 (10.83)	.013 (2.91)
Km to forest	−.938 (34.54)	−.938 (5.46)
Km to highway	.035 (1.24)	.035 (.140)
Tropical wet	3.145 (18.70)	3.145 (4.39)
Tropical moist		
Premontane wet	1.442 (33.69)	1.442 (4.08)
Premontane rain	−1.099 (23.38)	−1.099 (4.69)
Montane rain	−1.098 (22.14)	−1.099 (3.65)
Pseudo R^2		
Log Likelihood		

sample, the estimated standard errors obtained assuming simple random sampling will be underestimated. Adjustment factors can be calculated provided we are willing to identify each cluster precisely. To do so we again use the two alternative definitions of contiguity discussed before and proceed to reestimate a simple logit model with Huber correction which adjust upwardly all standard errors. The magnitude of the adjustment depends on the estimated intracluster correlations. The estimated effects and new t-values are displayed in Table A3. A glance at the table immediately reveals three important features: a) when the definition of contiguity is more demanding (k=20), the effects of landowners, landless peasants, demand for wood and NRR cease to be statistically significant. For all other variables the significance levels are maintained. Using a less demanding definition of contiguity (k=4) leaves the results from the original logistic model virtually unchanged.

The conclusions that one can draw from this exercise are mixed. First, we have estimated effects that are singularly sensitive to specification of spatial autocorrelation. Thus, the effects and associated standard errors of the variables landowners, landless peasants, and demand for wood are volatile and lead to contrasting conclusions about their influence on deforestation. It appears that landowners exert a positive effect on rates of deforestation whereas the size of the landless population operates as brake. Although this is not baffling at all since it supports the idea that the landless population lacks the capital and technical arsenal to successfully clear forests, confirmation of the results in alternative data sets is desirable.

Second, those variables that identify distances from road or cleared areas and the one associated with poverty are robust to model specification and should be used for hypothesis testing.

In summary, some of our original estimates are sensitive to spatial autocorrelation and cannot be used too recklessly to falsify hypotheses. In contrast, other estimates are distinctly robust and can be referred to quite liberally to confirm statements about the role of poverty, physical distance and, by implication, transportation and accessibility costs on deforestation.