Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize
Author(s): Kenneth M. Chomitz and David A. Gray
Source: *The World Bank Economic Review*, Vol. 10, No. 3 (Sep., 1996), pp. 487-512
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/3990055
Accessed: 09-09-2016 18:49 UTC

THE WORLD BANK ECONOMIC REVIEW, VOL. 10, NO. 3: 487-512

Roads, Land Use, and Deforestation: A Spatial Model Applied to Belize

Kenneth M. Chomitz and David A. Gray

Rural roads promote economic development, but they also facilitate deforestation. To explore this tradeoff, this article develops a spatially explicit model of land use and estimates probabilities of alternative land uses as a function of land characteristics and distance to market using a multinomial logit specification of this model. Controls are incorporated for the endogeneity of road placement.

The model is applied to data for southern Belize, an area experiencing rapid expansion of both subsistence and commercial agriculture, using geographic information system (GIS) techniques to select sample points at 1-kilometer intervals. Market access, land quality, and tenure status affect the probability of agricultural land use synergistically, having differential effects on the likelihood of commercial versus semisubsistence farming. The results suggest that road building in areas with agriculturally poor soils and low population densities may be a "lose-lose" proposition, causing habitat fragmentation and providing low economic returns.

The loss of tropical forests is a global concern because of its impact on biodiversity and climate. Roads are viewed as having precipitated much of this loss by opening forest areas to logging and agricultural conversion. This view poses a dilemma: road construction has traditionally been one of the most important tools for rural development (Creightney 1993) and, moreover, is thought to favor the rural poor (Lipton and Ravallion 1995). It is therefore important to quantify the impact of road building on both deforestation and development in order to assess the severity of the tradeoff between environmental preservation and economic growth.

Although planners are aware of the deleterious effect of roads on forests, they have no empirical guidance on the extent and nature of that damage. How far from the road do conversion effects extend? What kind of conversion is induced? Whom does it benefit? In this article we propose a model that addresses these questions. Following von Thünen (1966), we hypothesize that land

Kenneth M. Chomitz and David A. Gray are with the Policy Research Department at the World Bank. The authors are grateful to the Land Information Centre, Ministry of Natural Resources, Belize, for making the data available, and to the World Bank's Environment Department for providing geographic information system facilities. They thank Bruce King and Peter Orazem for beneficial discussions on an earlier draft, as well as William Hyde, David Wheeler, and an anonymous reviewer for helpful comments and suggestions. This project was supported by RPO #679-39 from the World Bank's Research Support Budget, and by the Project on the Social and Environmental Consequences of Growth Oriented Policies.

is devoted to the use that generates the highest potential rent. Roads play an important role in determining rent—and thus land use—by affecting agricultural output and input prices. But we hypothesize that the impact of roads will be strongly modulated by other factors affecting rent, including soil quality and distance from markets. If our hypothesis is correct—that road impacts are modulated by local conditions—then it may be possible to locate roads so as to spur development while minimizing induced deforestation.

Our goal is to develop a spatially explicit framework for testing this hypothesis. Spatially explicit models are appropriate for two reasons. First, they exploit rich spatial variation in variables of interest—variation that is obscured in aggregate data (for example, district-level means). Second, location matters. In general, we are interested not just in the physical extent of deforestation, but also the degree to which it affects critical habitats and watersheds.

The model we present uses spatially disaggregated data, controls for a wide variety of land and soil characteristics, employs multiple land-use categories, and is embedded in an economic framework. The model is estimated using 1989–92 data from Belize. Still mostly forested, Belize is of great interest for conservation because of its rich biodiversity and of its relatively large tracts of contiguous forest. Despite its small size, Belize exhibits many different deforestation processes, including encroachment by swidden agriculturalists and forest conversion to pasture, citrus groves, and large mechanized farms. In addition, Belize has superb documentation of land use and land characteristics, which facilitates this kind of study.

I. Issues and Analyses of Roads, Land Use, and Deforestation

Qualitatively, the impacts of roads on forests are clear. New roads offer market access for timber and agricultural products from previously remote areas. Roads also lower the costs of migration, land access, and land clearing for subsistence farmers. In sum, road construction into forested areas unambiguously increases the incentives to log those territories or convert them to other uses.

But the extent of deforestation hinges on the magnitude of those incentives, which we hypothesize to vary systematically over the landscape. Consider the following issues in regional and environmental planning:

- *Road siting and regional development.* Schneider (1995) and others have suggested that road building should be intensive rather than extensive. That is, road development should stress the creation of dense road networks around market centers rather than the extension of roads into areas with low population densities. Although this proposition seems to be reasonable, we lack information about the relative environmental costs and development benefits of the two strategies. Is road intensification a "win-win" strategy—that is, does it boost output and reduce environmental damage compared with an extensification strategy? What are its distributional implications?

The significance of these questions is underlined by the rapid expansion of road networks in the tropical world. In the 1980s, Brazil's paved road network grew from 87,000 kilometers to 161,500 kilometers, and Indonesia's from 56,500 kilometers to 116,500 kilometers (World Bank 1994). A recent review of Sub-Saharan Africa's rural roads concluded that "the present rural road network . . . needs to be increased up to tenfold if the full agricultural potential of the region is to be realized" (Riverson, Gaviria, and Thriscutt 1991).

- *Environmental impact assessment of forest and mining concessions.* According to some authors (Kummer and Turner 1994; Johnson and Cabarle 1993; Barbier and others 1994), logging's indirect impact on deforestation may be greater than the direct impacts of timber removal and collateral damage to standing stock. Forested regions are damaged additionally if logging roads and operations facilitate access by follow-on settlers, who convert the logged-over forest to pasture, permanent crops, or shifting cultivation. Hence, plans for the sustainable management of forest concessions must consider more than just silvicultural issues. Predictive models of follow-on settlement could be employed in environmental impact assessments of proposed logging and mining concessions that entail road building.
- *Conservation planning.* It is expensive to set up and maintain protected areas. Conservation planners have long recognized the need to establish an index of the threat of conversion to help prioritize candidate areas for protection. Such indexes have been constructed as ad hoc functions of population density or prior conversion, for example. But without behavioral grounding, these indexes may not be very accurate. Past deforestation rates, for instance, may be poor predictors of current rates if road networks have changed. Cross-sectional variation in population densities usually reflects differences in soil quality and may bear no relation to incentives for deforestation. Much more desirable would be a methodology that assesses actual deforestation incentives. This methodology could be used to design conservation policies that seek to alter these incentives.

In sum, quantitative models of land use and land-use change could be applied in a variety of environmental planning purposes. But to be useful, quantitative studies must meet several criteria. First, they must be based on spatially disaggregated data. Second, they must incorporate a wide range of land-use determinants, while recognizing that population distribution, road placement, and land-use change are jointly determined. Third, they must be based on an economic framework.

Few studies on developing countries meet all these criteria. A limited literature uses provincial- or district-level data to analyze patterns of deforestation. Alig (1986), Panayotou and Sungsuwan (1989), and Barbier and Burgess (1994) emphasize the role of prices in determining land use. The latter two studies include both prices and road accessibility in their analyses, but their estimates of net road impacts do not include the indirect impacts of roads through their effect on farmgate prices. Reis and Margulis (1991), Reis and Guzmán (1994), Southgate, Sierra, and Brown (1991), and Pfaff (1996) focus on the effects of roads, cattle, and population on deforestation. The latter two studies are noteworthy for attempting to deal with the endogeneity of population distribution.

Studies using large administrative areas as the unit of observation are inherently limited in their ability to discern how far across the landscape a road's impact stretches or to detect the impact of variations in soil quality on land use. Only a handful of studies use spatially disaggregated data. Liu, Iverson, and Brown (1993), writing on the Philippines, show that there is a strong inverse relationship between distance from the road network (as of 1941) and proportion of forest lost from 1934 to 1988. This suggestive bivariate relationship, however, may reflect the influence of other variables, such as soil fertility or market proximity. Three studies apply multivariate techniques to spatially disaggregated data. A study of Thai farmers by Fox and others (1994) uses a multinomial logit model to examine the determinants of crop choice among plots already converted to cultivation. Chomitz and Gray (1995) (a predecessor to this article) use an economic framework to derive a similar multinomial logit specification, extending it to include noncultivation as a land-use category. Nelson and Hellerstein (1995) use Chomitz and Gray's framework to examine the impact of road access and village access on land use in a region of central Mexico.

II. A SPATIAL MODEL OF LAND USE

Deforestation is just one aspect of a general model of land use. Below, we derive an estimatable, spatially explicit model of land use based on von Thünen (1966). Theoretical variants have been presented by von Amsberg (1994), Schneider (1995), and Hyde, Amacher, and Magrath (1993).

Following von Thünen, we suppose that there is a potential rent (farmgate value of output minus the cost of inputs) attached to each use of each plot of land. The model predicts, simply, that land will be devoted to the activity yielding the highest rent. In the classic example, farmers working near a city find that vegetables are more profitable to produce than grain (figure 1). Because vegetables are perishable, they are more expensive to transport than grain. Thus at some distance from the city grain becomes more profitable to produce than vegetables. And beyond a certain, even greater, distance the land is undisturbed under its original forest cover.

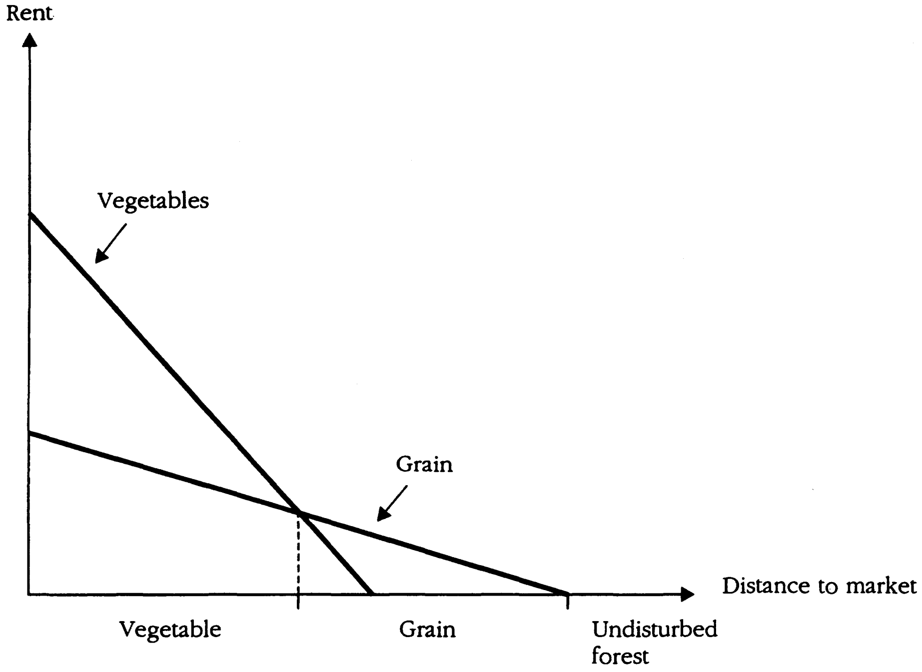
More formally, let P_{ik} be the price of the output of use k at point i, C_{ik} a vector of prices of inputs to k at that point, X_{ik} the optimal quantities of inputs for k per unit of land, and Q_{ik} the potential output of k at that point. The potential rent associated with devoting plot i to use or commodity k is[1]

$$R_{ik} = P_{ik}Q_{ik}(P_{ik}, C_{ik}) - C_{ik}X_{ik}(P_{ik}, C_{ik}). \tag{1}$$

---

[1] We are assuming a static framework. Dynamic issues are discussed later.

Figure 1. *The von Thünen Model Revisited*



Source: Based on von Thünen (1966).

Unfortunately, P, C, and Q are not only endogenous, they are also unobserved. We do, however, observe the determinants of price and of productivity and can therefore formulate a reduced-form model. (All the right-side variables are jointly determined with land use in a spatial equilibrium and are therefore endogenous. Hence, to estimate equation 1 directly, we would have to instrument these variables through auxiliary equations parallel to those used to derive the reduced form.)

Following von Thünen, we assume that spatial differentials in farmgate prices are related solely to differences in transport costs to major markets. For each commodity we can specify functions relating output and input prices to the distance D to market:[2]

$$P_{ik} = \exp(\gamma_{0k} + \gamma_{1k}D_i) \tag{2}$$

$$C_{ik} = \exp(\delta_{0k} + \delta_{1k}D_i).$$

Conventionally, output prices are assumed to decrease with distance ($\gamma_{1k} < 0$) and input costs to increase ($\delta_{1k} > 0$). The second assumption is very plausible for

---

[2] For greater flexibility, a polynomial in D could be specified. A linear specification (with truncation at zero) is more traditional. However, transport costs are not necessarily linear in distance. The exponential specification offers greater flexibility, but is used here primarily to permit derivation of an easily estimatable multinomial logit model.

bulk inputs such as fertilizer, but is less clear for labor inputs. There is little information on the spatial structure of wages in forested areas with low population densities. Scattered and anecdotal evidence suggests that the returns to labor are higher at the forest frontier than in areas closer to urban markets (Charras and Pain 1993; Schneider 1995). It is possible, however, in some places that labor costs (and thus, conceivably, total input costs) might decrease with distance from market.

The production function for use k, here expressed as output per unit of land, is assumed to be

$$Q_{ik} = S_{ik}X_{ik}^{\beta_k} \quad [0 < \beta_k < 1]. \tag{3}$$

The productivity factor S can be expressed as the product of agroclimatic and other variables, s_i, describing soil fertility and other land characteristics:

$$S_{ik} = \lambda_{0k} s_{1i}^{\lambda_{1k}} s_{2i}^{\lambda_{2k}} \ldots$$

From equation 3, we can derive the demand for X (suppressing subscripts for readability):

$$X = \left[\frac{C}{PS\beta}\right]^{1/[\beta-1]} \tag{4}$$

Substituting equations 3 and 4 into equation 1 (suppressing subscripts and simplifying), we have

$$R = PQ - CX = PSX^{\beta} - CX = X[PSX^{\beta-1} - C] = C^{\beta/(\beta-1)}[PS\beta]^{1/(1-\beta)}(1 - \beta)/\beta. \tag{5}$$

Together, equations 4 and 5 show that rent and input intensity increase as output prices increase, and decrease as input costs increase.

Substituting equations 2 and 3 into equation 5, taking logs, reintroducing subscripts, regrouping and simplifying coefficients, and introducing a stochastic error term, we have

$$\ln R_{ik} = \alpha_{0k} + \alpha_{1k}D_i + \alpha_{2k}\ln(s_{1i}) + \ldots + u_{ik} \equiv Z_iA_k + u_{ik} \tag{6}$$

where Z is the vector of independent variables, A is a vector of reduced form parameters, with $\alpha_{1k} = (\gamma_{1k} - \delta_{1k}\beta_{1k})/(1 - \beta_{1k})$, and $\alpha_{2k} = \lambda_{1k}/(1 - \beta_{1k})$.

We expect the coefficients on distance to be negative and those on productivity-enhancing land characteristics to be positive. The magnitude of those coefficients, however, will differ markedly from crop to crop. Bulky commodities should have a large (negative) coefficient on the distance measures. By contrast, farmers who produce mostly for subsistence may care relatively little about distance to market, yielding a coefficient with a small absolute value. It is theoretically possible for $\alpha_{1k} > 0$ in the special case of labor-intensive, high-value/weight commodities when wages decrease with distance to market.

To estimate the model, we assume that land is devoted to the highest-rent use: point i is devoted to use k if

$$R_{ik} > R_{ij}, \text{ for all uses } j \neq k. \tag{7}$$

If the disturbances u are Weibull distributed and uncorrelated across uses j, then equation 7 is equivalent to a multinomial logit model in which the probability that plot i is devoted to use k is

$$\text{Prob } (i \text{ devoted to } k) = \frac{\exp(Z_i A_k)}{\sum_j \exp(Z_i A_j)}. \tag{8}$$

The multinomial logit model allows us to estimate the coefficients in equation 6 provided that the coefficients of one use—for example, natural vegetation—are normalized to zero. With normalization, equation 8 can also be expressed as

$$\ln [\text{prob } (i \text{ devoted to } k)/\text{prob } (i \text{ devoted to comparison use})] = Z_i A_k. \tag{9}$$

The estimate $Z_i A_k$ can also be interpreted as the relative (but not absolute) rent for use k at location i.

To apply this highly stylized model, we must address a number of practical complications.

- *Road endogeneity.* The model assumes that the placement of the road network is exogenous to agricultural land use. In some applications this assumption is reasonable. For instance, some roads are installed for political reasons, to provide access to a mine site or to connect distant cities. In general, however, road construction and routing may be influenced by agricultural development considerations. If roads are preferentially routed through agriculturally suitable areas and if some aspects of suitability are not observed, then the model may overestimate the effect of distance from the road. A plot of land may be undeveloped not because it is far from the road—it may be far from the road because it is not suitable for development.

  In Chomitz and Gray (1995) we argued that this bias could be reduced by including a variety of soil quality indicators as land-use determinants. A major goal of this article is to apply an instrumental variables approach to control for road endogeneity. As instruments we use variables that are correlated with the measure of market distance, but that are arguably uncorrelated with unobserved determinants of land use (holding constant the observed determinants).
- *Price expectations.* Current land-use decisions depend not only on today's prices, but also on expectations of future prices. A classic example is using deforestation to assert land rights in an area in which land prices are expected to rise (Schneider 1995). Similarly, a transient decline in the price of a tree

crop will not lead to plantation abandonment. Ideally, then, the expected path of future prices should be included as an explanatory variable.

- *Reversibility of land use.* The model assumes that if cultivated land becomes uneconomic (because of a drop in crop prices, for example), it will revert to natural vegetation. This assertion is defensible, even in the short run, as long as "natural vegetation" is broadly defined. If "forest" is distinguished as a separate land use, reversibility would certainly not hold in the short run, since abandoned land takes years to return to forest. For many forest areas, reversibility might be a reasonable assumption in the context of a long-run, static equilibrium model. Large portions of Belize's forest, for instance, were recently leveled by hurricanes. Studies of the Amazonia, too, have shown that most abandoned plots quickly revert to forest (Moran and others 1994), though this is less likely for plots that have been intensively used or scraped by bulldozers (Nepstad, Uhl, and Serrão 1991). It is important to stress, however, that regrowth of forest cover does not necessarily imply maintenance of original biodiversity or carbon levels.
- *Tenure as a determinant of rent.* Equation 5 assumes that landowners will either adopt the highest-rent land use or rent or sell the land to someone else who will do so. But as Schneider (1995) and Hyde, Amacher, and Magrath (1993) have stressed, returns to different land uses depend strongly on tenure. On the frontier, where land rights are poorly defined and difficult to defend, it may not be profitable to invest in perennial crops. But with tenure security, perennials may represent the highest-value use of the land. Similarly, largeholders may refrain from renting out land to sharecroppers—even if sharecroppers enjoy higher returns—if land tenure might be jeopardized. Hence it is desirable to use the land's tenure status as an explanatory variable in equation 3.
- *Correlation of unmeasured influences across commodities.* The multinomial logit model requires that the unobserved effects on the rent for commodity k be independent of the unobserved effects for other commodities at the same point. This is a strong assumption: unmeasured aspects of soil fertility, for instance, may have similar effects on a variety of alternative crops. In future work we will apply a multinomial probit formulation in order to allow for correlation among unmeasured effects.

III. LAND USE IN BELIZE: CONTEXT AND RELEVANCE

Belize is a small country with about 200,000 inhabitants and about 22,000 square kilometers of land area. Less than 10 percent of the land area has been converted to agriculture or settlements, 65 percent is under broadleaf forest, and the remainder consists mostly of swamp, pine forest, and mangrove forest.

Despite its small size, Belize exemplifies many of the issues and circumstances surrounding deforestation. First, the population–forest area ratio, at 9.5 people per square kilometer, is of the same order of magnitude as that in a number of important forest regions, including Bolivia (14.5), Congo (11.2), and the Brazilian states of Amazonas (1.3) and Pará (4.6). Second, Belize is an important site for wildlife conservation because of its large tracts of contiguous forest and its high level of biodiversity. For example, despite its small size, Belize hosts 528 bird species compared with 650 for the entire United States (WRI 1994).

Third, the forests of Belize are facing increasing pressure from a wide variety of agricultural practices. In this article we focus on the southern part of the country, where rapid population growth among the Maya Indians, together with immigration from neighboring countries, has resulted in the expansion of traditional *milpa* (slash and burn) cultivation of maize for subsistence and rice as a cash crop. At the same time, large-scale farmers—many of them recent Mennonite immigrants—have cleared extensive areas for pasture and for mechanized production of food crops. Citrus cultivation increased rapidly over the 1980s in the study area. Interestingly, timber has been a mainstay of the Belizean economy for more than two centuries, but that industry has not been associated with deforestation because of the highly selective nature of logging.

IV. DATA AND ECONOMETRIC ISSUES

We estimate the model defined by equations 6 and 7 using data on a sample of land points. This information is derived from digitally coded maps using geographic information system (GIS) techniques. Sampling was performed by placing a 1-kilometer rectangular grid over the country south of the Western Highway, yielding 11,712 sample points. To visualize the data-extraction process, imagine that we stacked the data layers (maps) of interest. A pin pierces the stack at each sample point, and the mapped information for the point—slope, distance to road, soil quality—is recorded and collated. We excluded national parks and private reserves from our sample, because only five of the 2,147 sample points in these classes were used for agriculture.[3] We also excluded urban and water points. The data layers are briefly described below. For more detailed information, see Chomitz and Gray (1995).

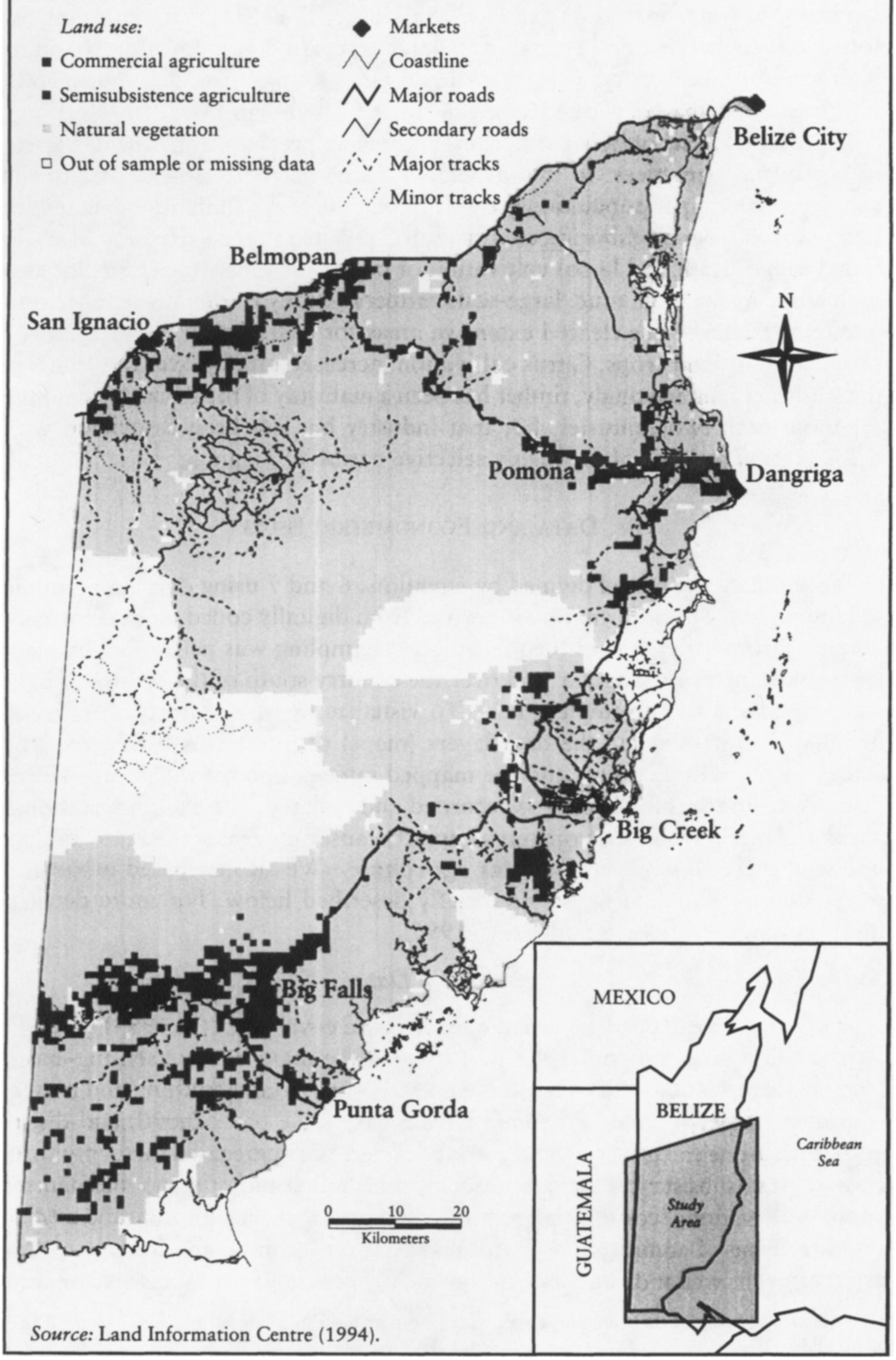
Land-Use Data

We derived the dependent variable from a land cover map (LIC 1994) based on SPOT satellite imagery from 1989 to 1992. Unlike many remote sensing–based land characterizations, this one employed extensive ground-truthing and may be considered highly reliable. The data have a base scale of 1:50,000, and distinguish thirty-one categories of land use/land cover. We aggregated these into three classes: "semisubsistence" agriculture, comprising milpa and other nonmechanized annual cultivation; "commercial farming," comprising mostly pasture and mechanized farming of annuals; and "natural vegetation," comprising forest, secondary regrowth, wetlands, and natural savanna. These land-use labels are for con-

---

[3] In effect, we assign an infinite negative coefficient to an indicator variable for these classes, which is, for all practical purposes, the maximum likelihood estimate.

Map 1. *Land Use in Southern Belize, 1989–92*



Source: Land Information Centre (1994).

venience only. Almost all of the "natural" vegetation in Belize has been modified by human action at some time in the past. Some of the areas labeled "semisubsistence" include market-oriented smallholders (map 1).

Land Systems and Land Tenure Data

The land systems data describe the soil's physical and chemical characteristics. These data are taken from a series of land resource assessments (King and others 1986, 1989, 1992), which were designed to yield planning information on the land's suitability for alternative crops.

The land resource assessments were based on a combination of aerial photography and field surveys. The methodology involved segmenting the landscape into about 10,000 internally homogenous microregions, based on agricultural potential as predicted from topography, soils, and vegetation. The microregions fall into 350 distinct classifications, called land subunits. These are similar to the land facets commonly associated with the land systems methodology. Each subunit is characterized by a set of physical and chemical descriptors, which in turn are used to assess the land's suitability for each of nineteen agricultural uses. It is important to stress that nutrient values are derived from field sites that are not used for agriculture. These values are then imputed to the same class of land subunits that are under cultivation.

Land in Belize falls into five broad tenure classes: private land, national land, Indian reserves, forest reserves, and protected areas. National land, which is held by the government but is available for lease by individuals, is believed to be more subject to encroachment by subsistence farmers than other tenure categories. The land systems and land tenure variables used in this analysis are as follows (for further details see Chomitz and Gray 1995; King and others 1993: 110–17):

Nitrogen: Soil nitrogen in percent
Slope: Slope in degrees
Available phosphorus: Available phosphorus in parts per million
pH: Soil pH
Wetness: An 8-point ordinal scale for drainage, ranging from 0 (well drained) to 7 (permanently wet)
Flood hazard: Dummy variable for flood hazard
Rainfall: Mean annual rainfall in meters
National land: Dummy for national land
Forest reserve: Dummy for forest reserve

Distance to Markets

Road network data are from a 1:50,000 topographic map series based on 1980s data, updated in certain areas to 1993. (It would have been preferable to recreate the road network as of 1989, the earliest date for the land-use data. We will attempt this in future work, using satellite imagery.) The measurement of market distance differs here from that used in Chomitz and Gray (1995). The earlier paper distinguished two components of distance: the straight-line distance from a sample point to the nearest point on the road network, and the subsequent on-road travel time to market (adjusted for road quality). This approach had the advantage of letting the estimation procedure yield relative weights for on- and off-road travel. Its disadvantage was that off-road distance was not adjusted for terrain—which varies substantially, from flat to mountainous to swampy.

In this study we compute an integrated distance measure by assigning impedance levels to different types of terrain. The impedances are judgmental estimates of the relative cost of transport. First- through fourth-class roads are assigned impedances of 1 through 4, respectively, representing the approximate inverse ratios of travel speed (60 miles through 15 miles per hour). Dry, level, roadless terrain is assigned an impedance of 100, based on a rough calculation of the relative cost of human-carried transport. More difficult terrain is assigned an impedance of up to 3,000 based on slope and wetness—implying that mountainous and swampy terrain is virtually impassable. To compute the distance from a point to a market, we gridded the landscape into 30-meter cells, assigned an impedance to each cell, and computed the route with the lowest cumulative impedance using standard iterative techniques.

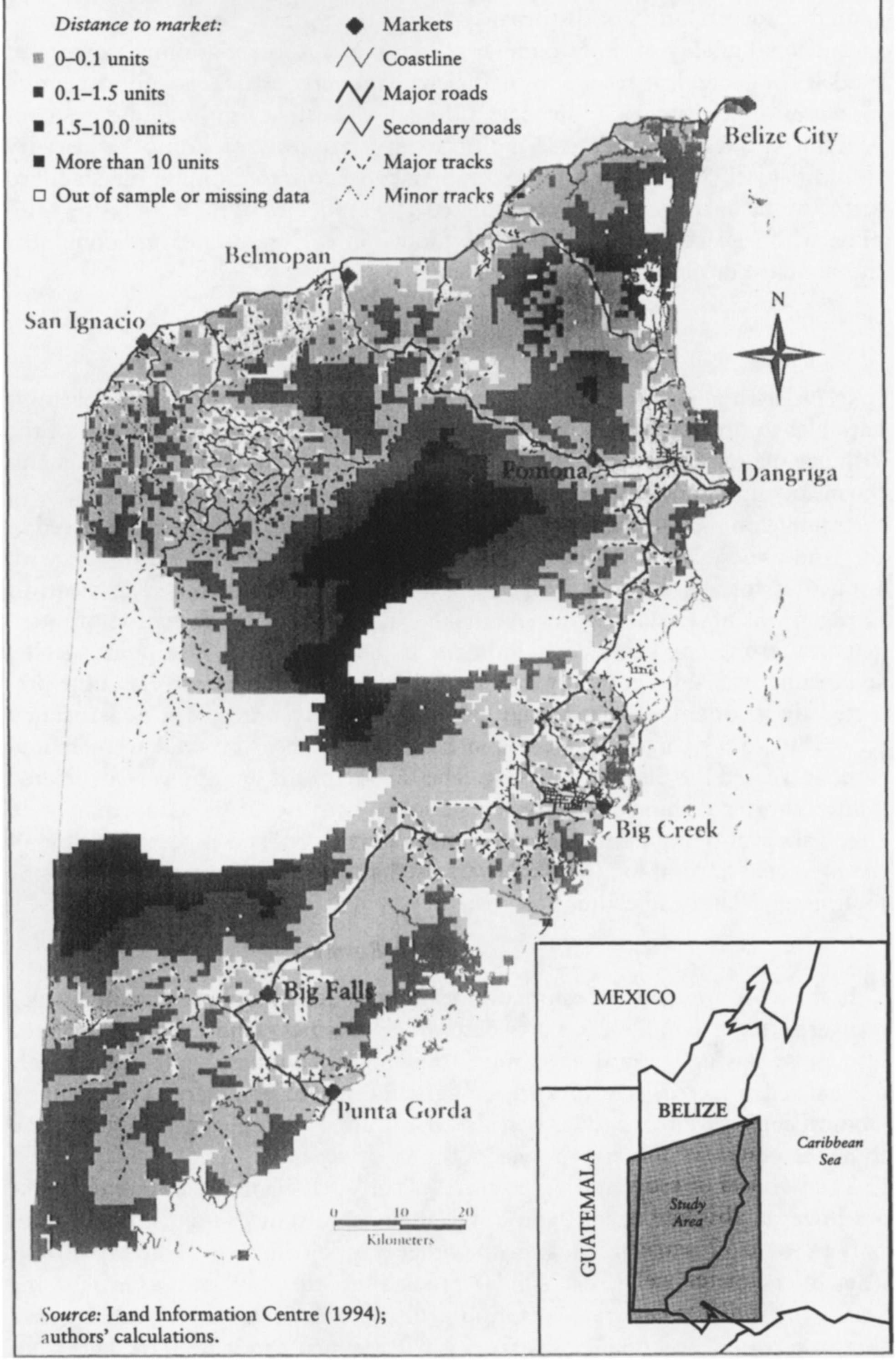
We computed distance to eight markets—Belize City, Belmopan, San Ignacio, Dangriga, Pomona, Big Creek/Independence, Big Fall, and Punta Gorda (map 2). Because of the low impedance assigned to on-road travel, distance to market is strongly related to distance from the nearest road.

Instruments for Distance

We calculated an accessibility measure in the same fashion as distance, only we calculated it as if there were no roads. That is, impedances for road cells were set to the impedance of the underlying terrain. This measure is strongly correlated with distance but is uncorrelated with unmeasured characteristics of a point's soil fertility, and thus constitutes an instrument for distance. The accessibility measure does, however, take market locations as exogenously given. This is a reasonable assumption in many cases—many of the towns in Belize started as logging or trading sites rather than as agricultural centers. The assumption is most problematic in the case of the citrus processing plant at Pomona, but even there the plant predated and influenced the massive citrus expansion of the 1980s. Additional instruments are provided by the mean and standard deviation of soil fertility measures within a 5-kilometer radius of the sample point. Holding constant its own soil fertility, a point is more likely to be served by roads if it is in a fertile area.

Spatial Autocorrelation

Land characteristics are likely to be characterized by spatial autocorrelation: places that are near to each other will tend to have similar soil types, rainfall, Map 2. *Integrated Cost Distance to Market, Southern Belize*



Source: Land Information Centre (1994); authors' calculations.

and so on. Some of these characteristics may not be observable. The resulting spatial autocorrelation of disturbance terms (the u_{ik} in equation 4) is the two-dimensional analog of more familiar patterns of autocorrelation in time-series models. In general, it results in inefficient parameter estimates and inaccurate measures of statistical significance, although consistency in limited dependent variable models requires our maintained hypothesis of homoskedasticity (McMillen 1992). We use a bootstrapping procedure to estimate the standard errors of the coefficients. We constructed 100 replicates of the data set by sampling with replacement, estimated the model on each replicate, and computed the standard deviations of the coefficients.

V. ESTIMATION RESULTS

In order to pursue the instrumental variables approach, we used exogenous variables to predict distance to market. For each market, the natural log of the distance measure was regressed on the natural log of the distance instrument, the mean and standard deviation of soil nitrogen within a 5-kilometer radius of the sample point, and the set of agroclimatic and tenure variables to be used in the land-use equation. All equations had an R^2 between 0.64 and 0.68, with t-statistics for the distance instrument ranging from 43 to 63. The mean of nitrogen was never statistically significant, but the standard deviation of nitrogen was strong and generally significant at the 0.001 level. The point-specific soil fertility variables generally had strong, highly significant effects in the predicted direction. A 0.1 percentage point increase in nitrogen was associated with a 24 to 33 percent reduction in the distance measure, and wetness and slope increased the distance measure. The flood hazard dummy was associated with a shorter distance measure, presumably because of its association with river valleys forming natural transportation corridors. The expected antilog of the predicted natural log (distance) was used as an explanatory variable in the multinomial land-use estimation.

Multinomial Logit Results

In the multinomial logit estimation using both actual and predicted distance measures (table 1; table 2 reports descriptive statistics) land use is aggregated into three classes: natural vegetation (the comparison class, with coefficients normalized to zero); semisubsistence agriculture; and commercial agriculture (comprising both citrus and large-scale agriculture). The relevant market is taken to be the nearest of the eight towns in the study area.

The two sets of estimates agree closely. The coefficients on market distance are lower in absolute magnitude using predicted distances—as expected—but only by 10 to 12 percent. Most of the other coefficients are reasonably robust. The two estimates yield very similar predictions: the correlation between the two predictions for natural vegetation probability is 0.96. This figure suggests that inclusion of soil quality variables is sufficient to eliminate most bias from road endogeneity. The following discussion will therefore employ the predictions based on actual distance. It is also noteworthy that bootstrap estimates of standard errors differ little from the standard estimates.

The estimates show that both semisubsistence and commercial agriculture become less attractive as distance to market increases, as expected. Commercial agriculture is much more sensitive to distance—a result that was not obvious a priori. Semisubsistence farmers market only a fraction of their production, so they might be expected to be insensitive to distance. On the other hand, large farms producing cattle or feed for integrated poultry production might also be insensitive to distance.[4] An alternative specification in which we assumed commercial farming to be oriented only to Belize City, rather than the nearest town, fit the data poorly. The predicted probability of commercial farming increased with distance to Belize City, reflecting the observed clustering of commercial farming around the smaller towns. We hypothesize that distance to the nearest market affects cropping at least partly through its effect on labor: farmers, both commercial and subsistence, are unwilling to live far from roads or towns.

Land and soil characteristics strongly affect the probability of agricultural use. Higher soil nitrogen boosts the probabilities of both types of agriculture, but has a far stronger effect on semisubsistence agriculture. Nitrogen is relatively more important for semisubsistence farmers, either because of differing crop mixes or (more likely) because of credit constraints in purchasing fertilizer. Both types of agriculture are encouraged by phosphorus and discouraged by an excessively low or high pH. These results are consistent with anthropological findings that traditional farmers use a variety of pedological and botanical cues to assess soil quality with considerable accuracy (see Carter 1969 on Maya farmers in Guatemala; Wilken 1987 on Mexican and Guatemalan farmers; and Moran 1993 on Brazilian farmers). The flood hazard dummy, a likely indicator of riverside location, strongly boosted the likelihood of semisubsistence farming but had a small, insignificant effect on the likelihood of commercial farming. Slope discourages commercial farming but encourages semisubsistence farming. National land has a low probability of commercial cultivation, but a high probability of semisubsistence cultivation, suggesting that these lands are subject to encroachment. Zoned forest reserves, however, have very low relative probabilities of agricultural use, all else equal.

Our principal conclusion is that market distance, land quality, and tenure have strong interactive effects on the likelihood and type of cultivation (figure 2). The top panel of figure 2 shows cultivation probabilities as a function of distance to market for the Cayo Floodplains land type, which is characterized

[4] Because citrus is far more sensitive to transport than other commercial crops, an alternative, four-land-use class model was also run. Citrus was distinguished from other types of commercial agriculture, and its market was assumed to be the processing plant at Pomona. As expected, citrus was found to be extremely sensitive to distance (with a coefficient of −8.0), but the remaining components of commercial agriculture were still more sensitive to distance than semisubsistence agriculture (with coefficients of −1.6 and −0.60). Soil quality coefficients for citrus were implausible and unstable because of the small number of citrus sample points. These results are available from the authors on request.

Table 1. *Multinomial Logit Estimates of Land Use, Belize*

	Using actual distances				*Using predicted distances*		
Variable	*Coefficient*	*Standard error*	*Standard error (bootstrap)*	*z-statistic (bootstrap)*	*Coefficient*	*Standard error*	*z-statistic*
Commercial agriculture							
Distance to market	−2.2553	0.2704	0.4543	−4.96	−1.9914	0.3007	−6.62
Soil nitrogen (percent)	5.1638	1.4776	1.1999	4.30	3.0062	1.4820	2.03
Slope (degrees)	−0.0173	0.0087	0.0092	−1.89	−0.0169	0.0098	−1.73
Available phosphorus (parts per million)	0.0429	0.0079	0.0068	6.32	0.0383	0.0077	5.00
Soil pH	1.4898	0.7917	0.7589	1.96	1.0133	0.7533	1.35
Soil pH squared	−0.1469	0.0739	0.0715	−2.05	−0.1041	0.0699	−1.49
Wetness[a]	−0.4606	0.1079	0.0949	−4.85	−0.4875	0.1057	−4.61
Wetness squared	0.0804	0.0179	0.0167	4.81	0.0775	0.0180	4.30
Flood hazard (dummy variable)	0.0901	0.1760	0.1672	0.54	−0.0736	0.1798	−0.41
Rainfall (annual mean in meters)	−0.3225	0.1262	0.1155	−2.79	−0.1437	0.1330	−1.08
National land (dummy)	−0.7175	0.1670	0.1691	−4.24	−0.6905	0.1649	−4.19
Forest reservation (dummy)	−3.2072	0.3917	0.4543	−7.06	−2.7627	0.3986	−6.93
Constant	−5.0524	2.1499	2.0037	−2.52	−3.7037	2.0347	−1.82

Semisubsistence agriculture							
Distance to market	−0.6002	0.0759	0.0921	−6.52	−0.5497	0.0980	−5.61
Soil nitrogen (percent)	16.8863	1.4455	1.3419	12.58	16.8703	1.4523	11.62
Slope (degrees)	0.0355	0.0073	0.0075	4.75	0.0335	0.0074	4.52
Available phosphorus (parts per million)	0.0421	0.0078	0.0094	4.49	0.0396	0.0077	5.17
Soil pH	2.3188	1.0439	1.1978	1.94	2.4429	1.0332	2.36
Soil pH squared	−0.1838	0.0974	0.1142	−1.61	−0.1935	0.0961	−2.01
Wetness[a]	1.0536	0.1497	0.1734	6.08	1.0430	0.1493	6.99
Wetness squared	−0.2188	0.0253	0.0308	−7.11	−0.2182	0.0254	−8.60
Flood hazard (dummy variable)	0.9188	0.1515	0.1611	5.70	0.8385	0.1505	5.57
Rainfall (annual mean in meters)	0.2749	0.0949	0.0989	2.78	0.3087	0.0990	3.11
National land (dummy)	0.7229	0.1089	0.1090	6.63	0.7623	0.1111	6.86
Forest reservation (dummy)	−1.8958	0.2210	0.2289	−8.28	−1.5904	0.2354	−6.76
Constant	−13.9614	2.7433	3.0562	−4.57	−14.4196	2.7233	−5.29
Log likelihood		−2,290			−2,360.4		
Number of observations		9,017			9,017		

Note: Natural vegetation is the comparison class (with coefficients normalized to zero). The relevant market is taken to be the nearest of the eight towns in the study area.

a. Wetness is calculated on an 8-point scale for drainage, ranging from 0 (well drained) to 7 (permanently wet).

Source: Authors' calculations.

Table 2. *Descriptive Statistics of Land Use, Belize*

Variable	*Mean*	*Standard deviation*	*Minimum*	*Maximum*
Distance to market	3.1884	4.7253	0.0004	24.7614
Nitrogen	0.1360	0.0498	0	0.2
Slope	16.3389	15.7382	0	40
Available phosphorus	5.2633	7.1391	2	31.5
pH	5.3180	0.9587	3.5	7.625
pH squared	29.2004	10.6575	12.25	58.1406
Wetness	1.7144	2.1941	0	7
Wetness squared	7.7530	12.6117	0	49
Flood hazard	0.4025	0.4904	0	1
Rainfall (meters)	2.3040	0.6186	1.025	4.046
Forest reservation	0.4628	0.4986	0	1
National land	0.2013	0.4010	0	1

Note: For description of variables, see section IV of the article.
Source: Authors' calculations.

by high phosphorus, moderate nitrogen, and low pH. This combination favors commercial agriculture. At the market there is a 34.4 percent chance that the land is converted to commercial agriculture use but only a 1.4 percent chance that it is used for semisubsistence cultivation. The remainder is under natural vegetation. The probability of commercial cultivation declines markedly with distance, however, so that at distance index = 1 (far from roads), commercial agriculture declines to 5.3 percent; it essentially vanishes if distance doubles.

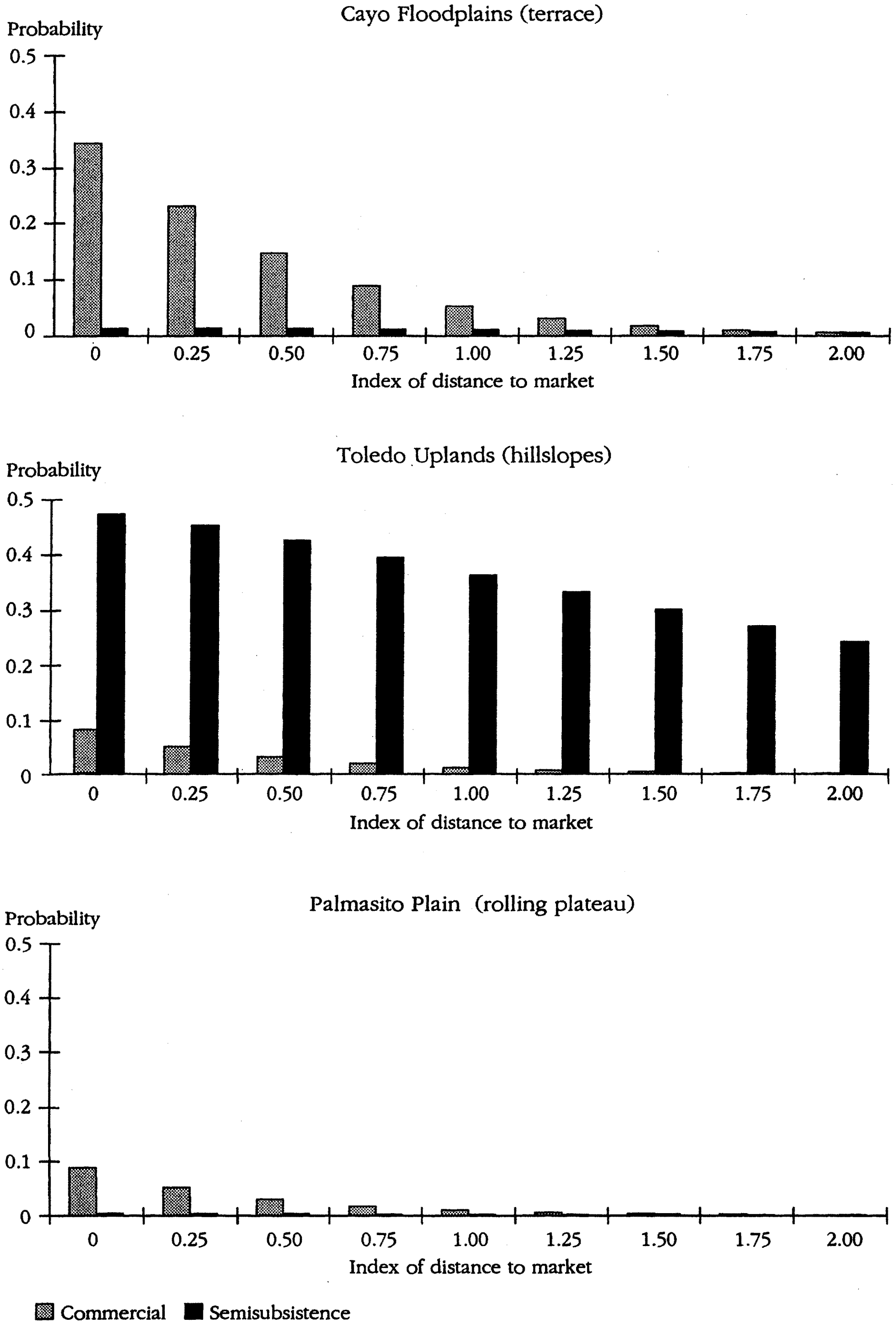
The middle panel of figure 2 predicts use of land typical of the Toledo Uplands, with high nitrogen and moderate slope, which favors semisubsistence agriculture. On this type of land, if the plot was adjacent to a market, the probability of semisubsistence cultivation would be 45.4 percent, compared with 5.0 percent for commercial agriculture. Note that the probability of semisubsistence use declines only gradually as distance to market increases, implying substantial cultivation even far from the road. This figure refers to current cultivation, and does not include the probability that land is in the fallow cycle of a shifting cultivation system.

The bottom panel of figure 2 shows predictions for land classified as Palmasito Plain. This poor-quality land has low nitrogen, low pH, and moderate slope. Even if adjacent to a market, such land would have only a 9.2 percent probability of being cultivated.

Predictive Ability of the Model

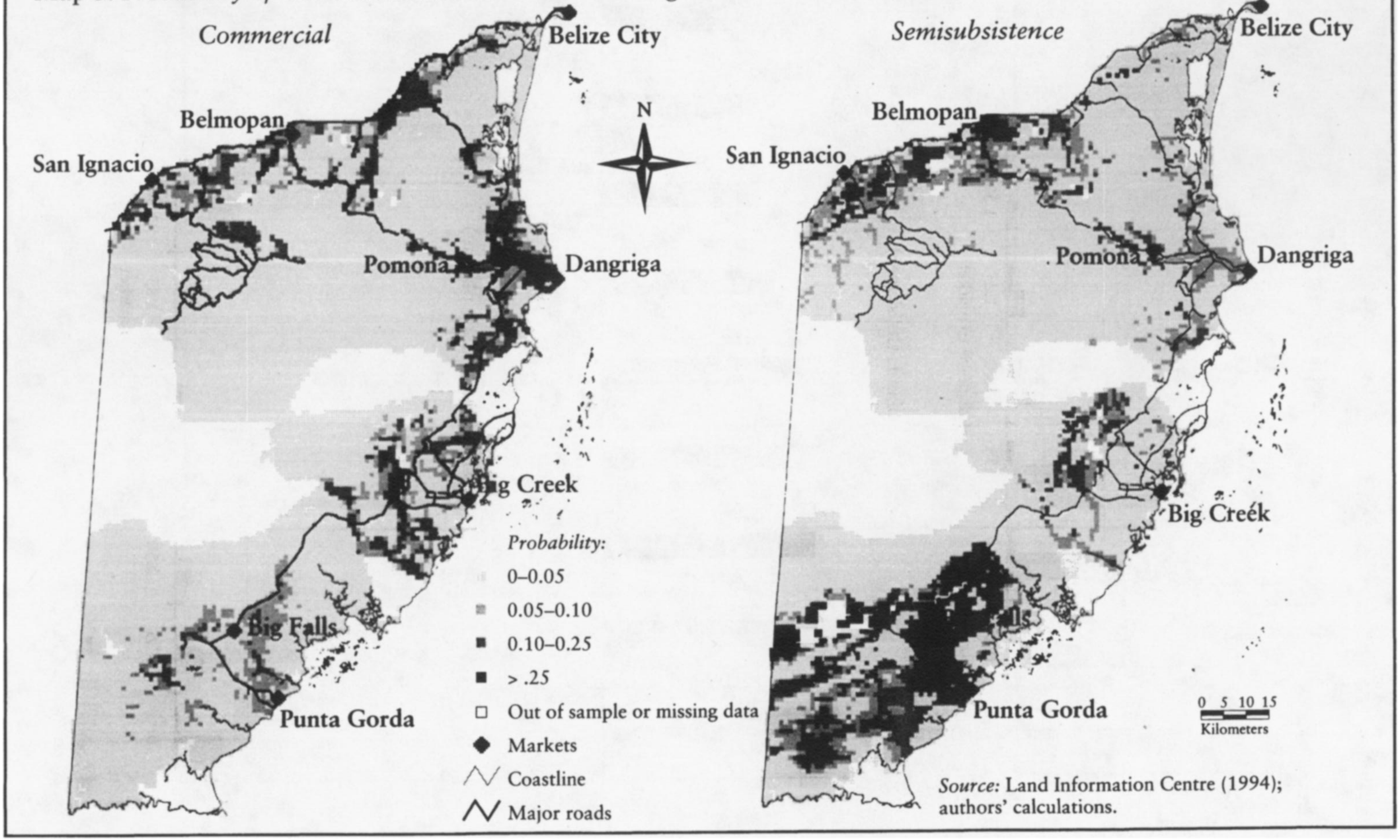
Qualitatively, the predicted probability of cultivation for commercial and semisubsistence agriculture (map 3) closely resembles actual land-use patterns shown in map 1. The model does not predict that particular points will have a very high probability of agricultural use. Rather, it predicts that entire areas are predisposed, with each individual point having a modest probability of cultivation. Is this fuzziness a vice or a virtue? It is a vice if it reflects the omission of

Figure 2. *Predicted Land-Use Probabilities, Belize*



Source: Authors' calculations.

Map 3. *Probability of Commercial and Semisubsistence Agricultural Land Use, Southern Belize*



Source: Land Information Centre (1994); authors' calculations.

important information that would differentiate the agricultural suitability of neighboring points. It is a virtue if neighboring points are in fact very similar in agricultural suitability. In that case it is truly a matter of chance which points are currently under cultivation. Thus the indication in map 3 of a broad southern area predisposed to semisubsistence agriculture may show the full range of the shifting cultivation system, including formerly cultivated areas now under forest fallow.

Assessing the model's predictive accuracy depends on two choices. The first is a rule for translating predicted land-use probabilities into discrete land-use category predictions. The conventional rule is to choose the category with the highest predicted probability. We argue, however, that even cultivation probabilities predicted to be low carry important information. We therefore suggest an alternative categorization rule: a point is predicted to be natural vegetation if the predicted probability of natural vegetation is greater than the sample mean probability (0.905). Otherwise, the prediction is assigned to the cultivation category with the highest predicted probability.

The second choice in assessing predictive accuracy is the relative tolerance of type-I and type-II errors. We can trivially achieve 90.5 percent overall predictive accuracy by predicting that all points are natural vegetation: all natural vegetation points would be correctly predicted, and all other points would be erroneously classified. Using our criterion, the model's overall accuracy is lower, at 73.6 percent, correctly classifying 74.1 percent of the natural vegetation points. However, it correctly classifies 68.8 percent of cultivated points and misclassifies only 12.5 percent of them as natural vegetation (table 3). These and subsequent predictive assessments exclude natural parks and private reserves. If we included these points and followed the rule that all such points are predicted to be under natural vegetation, the model's overall predictive performance would be much improved. Moreover, there may be meaning in these misclassifications. Some natural vegetation points are actually secondary growth or thicket, indicating recent cultivation and abandonment. Of these points, 48 percent were predicted to be under cultivation, compared with 25 percent of other natural vegetation points.

VI. POLICY APPLICATIONS AND FUTURE RESEARCH

The methodology described here can help planners assess the severity of environment-development trade-offs posed by road extension. We have shown that the impact of roads is highly sensitive to soil quality and to tenure regulations. Some new roads will favor forest clearing for commercial crops, while others will stimulate the spread of shifting cultivation. In the latter cases the planner may be faced with difficult choices—forest preservation, development, poverty alleviation—depending on the characteristics of the farming systems in question. But road extensions into many areas with poor or mediocre soil—which dominate tropical areas—could constitute a lose-lose strategy. The economic return to such roads might be low, given the modest conversion rates that they

Table 3. *Predicted and Actual Land Use, Belize*

	Predicted land use			
Statistic	*Commercial agriculture*	*Semi-subsistence*	*Natural vegetation*	*Actual land use*
Commercial agriculture				
Number of sample points	185	85	39	309
Percentage of actual sample points	60	28	13	100
Percentage of total predicted sample points	17	5	1	3
Semisubsistence agriculture				
Number of sample points	76	407	69	552
Percentage of actual sample points	14	74	13	100
Percentage of total predicted sample points	7	23	1	6
Natural vegetation				
Number of sample points	857	1,258	6,041	8,156
Percentage of actual sample points	11	15	74	100
Percentage of total predicted sample points	77	72	98	90
Total				
Number of sample points	1,118	1,750	6,149	9,017
Percentage of actual sample points	12	19	68	100
Percentage of total predicted sample points	100	100	100	100

Note: This assessment excludes natural parks and private reserves.
Source: Authors' calculations.

will stimulate. And even relatively small amounts of conversion along the road corridor can result in habitat fragmentation, threatening the viability of some populations. This is of particular concern because recent ecological studies suggest that tropical forests based on poor soils have higher levels of biodiversity than those based on good soils (Huston 1994). In addition, road access would expose the forest to various forms of degradation, such as overextraction of mahogany or poaching of birds.

At first glance, these results appear to be obvious. We believe, however, that the exercise is useful on three counts. First, it was not clear a priori that an economic model would be successful in describing the landscape. Second, the results for semisubsistence cultivation, though plausible, were not obvious a priori. Many observers believe that subsistence farmers are insensitive to soil conditions and will thus colonize along any available roadway. To the contrary, we found that in the low population-density context of Belize, subsistence farmers shun poor-quality land and are moderately sensitive to market access. This finding suggests that the hypothesis that logging causes damage primarily by inducing follow-on migration along logging roads does not necessarily apply to remote, low population-density areas with poor soils. Finally, the methodology presented yields quantitative results. This information can in turn be used to help assess environmental impacts and economic benefits.

The model has shortcomings that we hope to correct in future work. First, the assumptions underlying the multinomial logit specification are strong—they will be relaxed by using multinomial probit or nested logit specifications. Second, the model does not explicitly incorporate prices or price formation. Once prices are explicitly built into the model, the model can be used to examine the environmental impact of changes in agricultural, trade, or macroeconomic policies through their effect on output and input prices. Finally, the spatial framework presented here can be used to consider a broader class of environmental impacts, such as pollution from agricultural runoff.

We stress that this article is only a sketch of an analytic strategy. A thorough analysis would involve calculation of the impacts of particular road-siting alternatives. It would also allow for general equilibrium effects: a substantial increase in cropped area would boost wages and reduce the prices of domestically consumed agricultural products, changing the coefficients embodied in the model. Finally, it would take into account distributional effects (across income groups or regions) of altered cultivation patterns.

Our methodology also has applications to conservation planning, because there is a need for techniques that predict threats to protected areas. The tools presented here emphasize the use of soil maps for assessing the spatial patterns of threat, and also help to predict the nature of that threat—information that could be useful in designing integrated development and conservation programs.

These techniques can also be used to gauge the effectiveness of habitat protection. In Belize we observe very little cultivation in national parks and private reserves (about 0.5 percent). Is this because these areas are effectively policed, or is it because they are remote or otherwise unattractive for cultivation? To address this question, we use the estimated coefficients to predict the extent of cultivation to which these areas would be subject if they were not protected. Taking the expectation of predicted probabilities, we estimated that 5 percent of the national park area in southern Belize would be under current cultivation, suggesting that habitat protection has been effective in Belize.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Alig, R. J. 1986. "Econometric Analysis of Forest Acreage Trends in the Southeast." *Forest Science* 32(1):119–34.

Barbier, Edward B., Nancy Bockstael, Joanne Burgess, and Ivar Strand. 1994. "The Timber Trade and Tropical Deforestation in Indonesia." In Katrina Brown and David W. Pearce, eds., *The Causes of Tropical Deforestation.* London: UCL Press.

Barbier, Edward B., and Joanne C. Burgess. 1994. "Economic Analysis of Deforestation in Mexico." Discussion Paper 9406. University of York, Department of Environmental Economics and Environmental Management, York, U.K. Processed.

Carter, William. 1969. *New Lands and Old Traditions: Kekchi Cultivators in the Guatemalan Lowlands.* Latin American Monographs, second series. Gainesville, Fla.: University of Florida Press.

Charras, Muriel, and Marc Pain, eds. 1993. *Spontaneous Settlements in Indonesia.* Bondy, France: ORSTOM.

Chomitz, Kenneth M., and David A. Gray. 1995. "Roads, Lands, Markets, and Deforestation: A Spatial Model of Land Use in Belize." Policy Research Working Paper 1444. World Bank, Policy Research Department, Washington, D.C. Processed.

Creightney, Cavelle D. 1993. *Transport and Economic Performance: A Survey of Developing Countries.* World Bank Technical Paper 232. Washington, D.C.

Fox, Jefferson, Rebekah Kanter, Sanay Yarnasarn, Methi Ekasingh, and Royce Jones. 1994. "Farmer Decision Making and Spatial Variables in Northern Thailand." *Environmental Management* 18(3):391–99.

Huston, Michael A. 1994. *Biological Diversity: The Coexistence of Species on Changing Landscapes.* Cambridge, U.K.: Cambridge University Press.

Hyde, William F., Gregory S. Amacher, and William Magrath. 1993. "Deforestation, Scarce Forest Resources, and Forest Land Use: Theory, Empirical Evidence, and Policy Implications." Draft revision of paper presented at the Fourth Annual Conference of the International Association for the Study of Common Property, Manila, June 15–19. Processed.

Johnson, Nels, and Bruce Cabarle. 1993. *Surviving the Cut: Natural Forest Management in the Humid Tropics.* Washington, D.C.: World Resources Institute.

King, R. B., I. C. Baillie, P. G. Bisset, R. J. Grimble, M. S. Johnson, and G. L. Silva. 1986. *Land Resource Survey of Toledo District, Belize.* Tolworth, U.K.: Land Resource Development Centre.

King, R. B., I. C. Baillie, J. R. Dunsmore, R. J. Grimble, M. S. Johnson, and A. C. S. Wright. 1989. *Land Resource Assessment of Stann Creek District, Belize.* Chatham, U.K.: Overseas Development Natural Resources Institute.

King, R. B., I. C. Baillie, T. M. B. Abell, J. R. Dunsmore, D. A. Gray, J. H. Pratt, H. R. Versey, A. C. S. Wright, and S. A. Zisman. 1992. *Land Resource Assessment of Northern Belize.* Chatham, U.K.: Overseas Development Natural Resources Institute.

King, R. B., J. H. Pratt, M. P. Warner, and S. A. Zisman. 1993. *Agricultural Development in Belize.* Bulletin 48. Chatham, U.K.: Overseas Development Natural Resources Institute.

Kummer, David M., and B. L. Turner II. 1994. "The Human Causes of Deforestation in Southeast Asia." *Bioscience* 44(5):323–28.

Land Information Centre (LIC). 1994. "The Land Use of Belize, 1989/92." Technical Report. Ministry of Natural Resources, Land Information Centre, Belmopan, Belize. Processed.

Lipton, Michael, and Martin Ravallion. 1995. "Poverty and Policy." In J. Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics* III. New York: Elsevier.

Liu, Dawning S., Louis R. Iverson, and Sandra Brown. 1993. "Rates and Patterns of Deforestation in the Philippines: Application of Geographic Information System Analysis." *Forest Ecology and Management* 57:1–16.

McMillen, Daniel P. 1992. "Probit with Spatial Autocorrelation." *Journal of Regional Science* 32(3):335–48.

Moran, Emilio F. 1993. *Through Amazonian Eyes: The Human Ecology of Amazonian Populations*. Iowa City: University of Iowa Press.

Moran, Emilio F., Eduardo Brondizio, Paul Mausel, and You Wu. 1994. "Integrating Amazonian Vegetation, Land-Use, and Satellite Data." *BioScience* 44(5):329–38.

Nelson, Gerald C., and Daniel Hellerstein. 1995. "Do Roads Cause Deforestation? Using Satellite Images in Econometric Analysis of Land Use." Staff Paper 95 E-488. University of Illinois, Department of Agricultural Economics, Urbana-Champaign, Ill. Processed.

Nepstad, Daniel C., Christopher Uhl, and Emmanuel A. S. Serrão. 1991. "Recuperation of a Degraded Amazonian Landscape: Forest Recovery and Agricultural Restoration." *Ambio* 20(6):248–55.

Panayotou, Theodore, and Somthawin Sungsuwan. 1989. "An Econometric Study of the Causes of Tropical Deforestation: The Case of Northeast Thailand." Discussion Paper 284. Harvard Institute of International Development, Cambridge, Mass. Processed.

Pfaff, Alexander S. P. 1996. "What Drives Deforestation in the Brazilian Amazon? Evidence from Satellite and Socioeconomic Data." Draft. Columbia University, Department of Economics, New York. Processed.

Reis, Eustáquio, and Rolando Guzmán. 1994. "An Econometric Model of Amazonian Deforestation." In Katrina Brown and David W. Pearce, eds., *The Causes of Tropical Deforestation*. London: UCL Press.

Reis, Eustáquio, and Sergio Margulis. 1991. "Options for Slowing Amazon Jungle Clearing." In Rudiger Dornbusch and James M. Poterba, eds., *Global Warming: Economic Policy Responses*. Cambridge, Mass.: MIT Press.

Riverson, John, Juan Gaviria, and Sydney Thriscutt. 1991. *Rural Roads in Sub-Saharan Africa: Lessons from World Bank Experience*. World Bank Technical Paper 141. Washington, D.C.

Schneider, Robert R. 1995. *Government and the Economy on the Amazon Frontier*. World Bank Environment Paper 11. Washington, D.C.

Southgate, Douglas, Rodrigo Sierra, and Lawrence Brown. 1991. "The Causes of Tropical Deforestation in Ecuador: A Statistical Analysis." *World Development* 19(9):1145–51.

von Amsberg, Joachim. 1994. "Economic Parameters of Deforestation." Policy Research Department Working Paper 1350. World Bank, Policy Research Department, Washington, D.C. Processed.

von Thünen, J. H. 1966. "Der Isolierte Staat in Beziehung der Landwirtschaft und Nationalökonomie." In Peter Hall, ed., *Von Thünen's Isolated State*. Oxford, U.K.: Pergamon Press.

Wilken, Gene C. 1987. *Good Farmers: Traditional Agricultural Resource Management in Mexico and Central America.* Berkeley, Calif.: University of California Press.

World Bank. 1994. *World Development Report 1994: Infrastructure for Development.* New York: Oxford University Press.

WRI (World Resources Institute). 1994. *World Resources 1994–95.* New York: Oxford University Press.