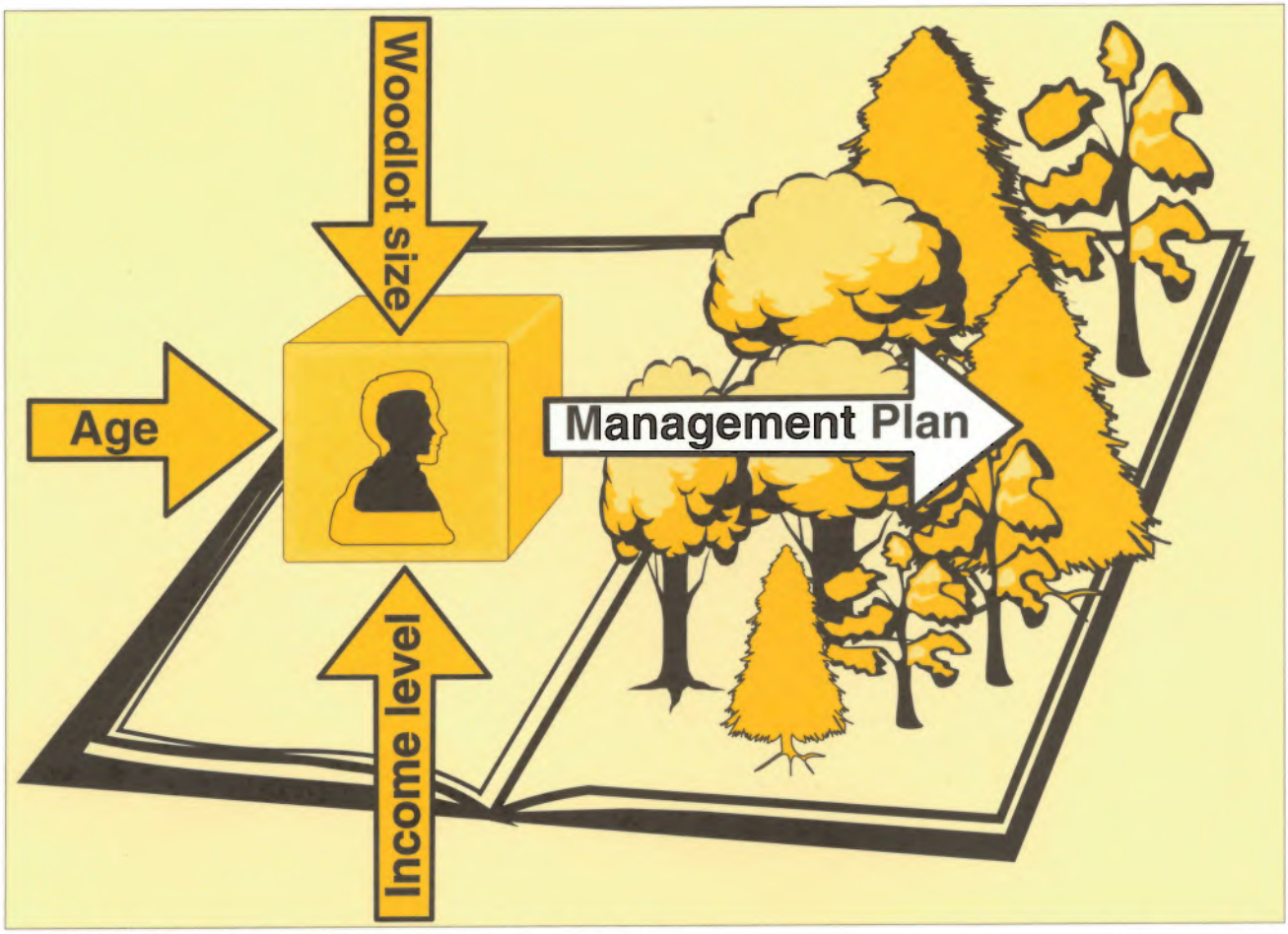
Predicting forest management on private woodlots in Carleton and Victoria counties using logit analysis

By Derek MacFarlane
Forest Economics Group

Canadian Forest Service - Maritimes Region
Information Report M-X-189E



Natural Resources Canada
Ressources naturelles Canada
Canada

PREDICTING FOREST MANAGEMENT ON PRIVATE WOODLOTS IN CARLETON AND VICTORIA COUNTIES IN NEW BRUNSWICK USING LOGIT ANALYSIS

Derek MacFarlane
Forest Economics Group

M - X - 189E

Canadian Forest Service - Maritimes Region
P. O. Box 4000, Fredericton, N.B. E3B 5P7

1994

ISSN: 1195-3799
ISBN: 0-662-22412-4
Catalogue No.: Fo46-19/189E

Additional copies of this publication are available in limited quantities at no charge from:

Canadian Forest Service - Maritimes Region
P.O. Box 4000
Fredericton, New Brunswick
Canada E3B 5P7
(506) 452-3500
Fax (506) 452-3525

Photocopies or microfiches of this publication may also be purchased from:

Micromedia Ltd.
Place-du-Portage
165, rue Hôtel-de-Ville
Hull (Québec)
J8X 3X2

Une copie française de ce rapport est disponible sur demande.

Canadian Cataloguing in Publication Data

MacFarlane, Derek David, 1958-

Predicting forest management on private woodlots in Carleton and Victoria counties in New Brunswick using logit analysis

(Information report ; M-X-189E)
Issued also in French under title: La prévision de l'aménagement forestier des boisés privés au sein des comtés de Carleton et Victoria au Nouveau-Brunswick au moyen de l'analyse logit.
Includes bibliographical references.
ISBN 0-662-22412-4
Cat. no. Fo46-19/189E

1. Woodlots — New Brunswick — Management.
2. Forest Management — New Brunswick.
I. Canadian Forest Service - Maritimes Region.
II. Title.
III. Series: Information report (Canadian Forest Service - Maritimes Region) ; M-X-189E.

SD387.W6M32 1994 634.9'2'097151 C94-980269-7

Abstract

This paper uses a logit model to predict the liklihood of a landowner engaging in forest management based on his or her characteristics. Information from a 1990 survey conducted by one of the New Brunswick forest products marketing boards, the Carleton-Victoria Wood Producers Association (CVWPA), was used in the analysis. The original 17-variable model was reduced to five variables that had the same degree of accuracy as the original 17-variable model. Compared to a similar province-wide analysis, this study indicates that regional differences exist when it comes to the type of owner likely to engage in management. In the CVWPA area, the decision to manage increases if total area owned increases and total income rises. The probability decreases if education level is grade 12 or less and as income derived from wood increases.

Keywords: forest management, logit analysis, New Brunswick

Résumé

Ce document porte sur l'application d'un modèle logit à la prévision de la probabilité qu'un propriétaire foncier se livre à des activités d'aménagement forestier en fonction de certaines de ses caractéristiques. L'information tirée d'une enquête réalisée en 1990 par l'un des offices de commercialisation des produits de la forêt du Nouveau-Brunswick, le Carleton-Victoria Wood Producers Association (CVWPA), a servi aux fins de cette analyse. Le modèle original à 17 variables a été réduit à 5 variables possédant le même degré de précision. En comparaison d'une analyse similaire réalisée à l'échelle de la province, l'étude révèle l'existence de différences régionales lorsqu'on tient compte du type de propriétaires susceptibles de se livrer à des activités d'aménagement. Au niveau du secteur CVWPA, la probabilité que les propriétaires décident d'aménager augmente lorsqu'augmentent la superficie des propriétés et le revenu total. Cette probabilité diminue d'autre part lorsque le niveau d'instruction est inférieur à une douzième année ou lorsque augmentent les revenus tirés du bois.

Mots clés : aménagement forestier, analyse logit, Nouveau-Brunswick

Table of Contents

List of Tables

Introduction

The forest industry in New Brunswick (N.B.) is very important to the provincial economy. In 1986, almost 35% of N.B.'s gross domestic product (GDP) in the primary sector resulted from logging activity (Statistics Canada, 1990). In the same year, the forest industry represented almost 39% of the total GDP in the manufacturing sector.

In N.B., the majority of productive forest land is owned by either the province, industry or private woodlot owners (50%, 20%, and 30%, respectively).

The forest industry relies on a sustainable supply of raw material. Industrial requirements cannot be met from harvesting on one or even two of these tenures. All three tenures are part of the sustainable supply equation. Wood supply analysis in N.B. (based on current consumption and technology) predicts that in 15-20 years there will be a shortfall in wood. One strategy that will help overcome this shortfall is a forest management program on all tenures. Private woodlots are part of this strategy, but with 35,000 landowners, it becomes very difficult to predict whether the forest management targets established to overcome the shortfall will indeed be met.

Clements and Jamnick (1989) used a logit model to predict the likelihood of a landowner engaging in forest management based on his or her characteristics. The data used in their study was from a 1983 survey of all woodlot owners in N.B. Almost 9,000 respondents were used in their analysis. Their study was able to reduce a 22-variable model (full) to a seven-variable model that had a predictive capability and significance level nearly identical to the full model and had coefficients with the expected signs.

This paper will use information from a 1990 survey conducted by the Carleton-Victoria Wood Producers Association (CVWPA). The survey questions were designed so that a direct comparison could be made with the Clements and Jamnick (1989) study. A logit model will be developed from the results of this survey and compared to the Clements and Jamnick (1989) results.

The Logit Model

The logit model belongs to the group of nonlinear models in which the dependent variable involves two (or more) qualitative choices (Pindyck and Rubinfeld, 1991). This "binary choice" model assumes that individuals are faced with a choice between two alternatives and that their choice depends on their characteristics. Therefore, the objective of the model developed in this paper will be to determine the probability that an individual (woodlot owner) with a specific set of attributes will make a choice between managing or not managing his/her woodlot. This type of analysis was used by Gramann *et al.* (1985) with relative success. Their study was based on owner beliefs and the probability of carrying out forest management based on these beliefs.

The logit model is estimated by a method called Maximum Likelihood Estimation (MLE). Ordinary regression models are estimated by the method of Ordinary Least Squares (OLS). The conceptual difference between OLS and MLE is that OLS is concerned with picking parameter estimates that yield the smallest sum of squared errors in the fit between the model and data, while MLE is concerned with picking parameter estimates that imply the highest probability or likelihood of having obtained the observed sample (Aldrich and Nelson, 1984).

The assumptions of the logit model are as follows:

(i) The dependent variable, Y, is binary.

(ii) The relationship between the dependent and independent variables can be estimated by:

$$log \frac{Pi}{1-Pi} = F(Zi) \qquad [1]$$

where, Pi = probability that an individual will make a certain choice given Xi (independent variables)

Zi = right-hand side (*i.e.*, $a + \beta x$).

(iii) The data are generated from a random sample size N. This requires that the observations on Y be statistically independent of each other, ruling out serial correlation.

(iv) There is no exact or near linear dependency among the independent variables (*i.e.,* no multicollinearity).

Aldrich and Nelson (1984) indicate that the exact properties of MLE (lack of bias, efficiency, normality) cannot be established. Nevertheless, these properties hold, approximately, with the quality of approximation improving as the sample size grows.

The logit model in this report will be subjected to a number of tests formally described in Aldrich and Nelson (1984) and Pindyck and Rubinfeld (1991).

Methods

In 1990, the CVWPA conducted a survey of 1,200 landowners in Carleton and Victoria Counties. There were 404 usable responses, representing 33.7% of the total. Forty questions were asked which covered various owner and ownership characteristics. Also, each owner was asked if any forest management activity had been undertaken in the past 5 years. Similar to the Clements and Jamnick (1989) study, only those management activities that would influence timber supply were considered. This excluded activity related to Christmas trees and maple stands.

If at least one of the activities related to timber supply was chosen, the dependent variable MANAGE was coded 1. If there was no management activity, MANAGE was coded 0. Also, the question relating to harvesting forest products was restricted to pulpwood, sawlogs, veneer logs, posts, poles, pilings or rails, rather than including Christmas trees, maple syrup products, and firewood for the same reason as above.

The entire data set (404) was used to estimate the model. Missing data were coded -1. The owner and ownership characteristics used in the model totalled 17 compared to the 22 used by Clements and Jamnick (1989). These characteristics were hypothesized to influence an owner's decision to carry out forest management and are shown in Table 1. A logit model was estimated using all 17 variables (Doan, 1989).

The results for Model I are summarized below. The standard errors are in parentheses below the estimated coefficient and * indicates that the estimate is significant at the 5% level.

$$\log \frac{P(\text{manage})}{1-P(\text{manage})} = \quad [2]$$

- 2.2339 + 0.2511 AGE + 0.0343 COLLEGE
(2.0523)* (0.2450) (0.6693)

- 0.1974 DIST + 0.0027 FAM - 0.49 FARM
(0.1932) (0.0050) (0.6464)

+ 0.6773 HAVESOLD - 0.0053 INSTRUCT - 0.1911 OBTAIN
(0.5293) (0.1001) (0.3992)

- 0.5765 PROFESS - 1.3631 RETIRE - 1.4126 SCHOOL12
(0.8137) (0.7940) (0.5715)*

+ 0.0004 TOTACR - 0.2688 TOTINC - 2.1663 TRADESC
(0.0005) (0.2154) (0.8524)*

- 0.4845 WOODINC - 0.114 YOBTAIN
(0.3038) (0.1973)

The model correctly classified 91.4% of the data used in the model construction. Besides the intercept, TRADESC and SCHOOL12 were the only variables that produced coefficients significantly different from zero. There were two measures for "goodness of fit" applied. The first uses a likelihood ratio index described by Pindyck and Rubinfeld (1991).

$$p = 1 - \frac{L(\beta^*)}{L(0)} \quad [3]$$

where, $0 \leq p \leq 1$

and

$L(\beta^*)$ = value of the log likelihood with the full model

$L(0)$ = value of the log likelihood with parameters = 0

$$\text{Therefore, } p = 1 - \frac{(-83.82)}{(-119.05)} = 0.30 \quad [4]$$

Table 1. Variable names, expected signs, coding and definitions for the independent variables used in the model.

Variable	Expected Sign[1]	Definition and Coding
Age	-	Age of the woodlot owner? Coded 1<25; 2=35-34; 3=35-44; 4=45-54; 5=55-64; 6=>65
College	+	College or University education? Coded 0=no; 1=yes
Dist	-	Distance owner lives from woodlot in miles? Coded 1>5; 2=5-24; 3=25-44; 4=50-99; 5=≥100
Fam	+	Number of years inherited land has been in the family? Coded "number" of years
Farm	+	Main occupation: farmer? Coded 0=no; 1=yes
Havesold	+	Forest products sold in past 5 years? Coded 0=no; 1=yes
Instruct	+	Have had forestry or logging instruction? Coded 0=no; 1-8 depending on type of instruction
Obtain	+	How was your woodlot obtained? Coded 1=inherited; 2=bought; 3=inherited and bought; 4=other
Profess	+	Main occupation: Professional? Coded 0=no; 1=yes
Retire	-	Woodlot owner retired? Coded 0=no; 1=yes
School12	?	High school education or less? Coded 0=no; 1=yes
Totac	+	Total number of acres owned? Coded "number" of acres
Totinc	+	Average annual income from all sources ($)? Coded 1<4999; 2=5000-14,999; 3=15,000-24,999; 4=25,000-34,999; 5=35,000-44,999; 6>45,000
Tradecs	+	Education includes trade, technical, or commercial school? Coded 0=no; 1=yes
Woodinc	+	Percentage of total income earned from the woodlot? Coded 1=0; 2=1-10; 3=11-25; 4=26-50; 5=51-70; 6=76-100.
Yobtain	+	Year first parcel of forestland was obtained? Coded 1=<1940; 2=1940-1949; 3=1950-1959; 4=1960-1969; 5=1970-1979; 6>1980

[1]Expected sign, for example for variable "age" means that the older the owner, the less likely the decision will be to manage (*i.e.* "-" sign). Similarly for the variable "College", the more education the owner has, the more likely is the decision to manage (*i.e.*, "+" sign).

The values for p will range from 0 to 1. Pindyck and Rubinfeld (1991) warn that with this type of model (*i.e.*, logit), the p value (like R^2) will not be close to 1 and that the value is difficult to interpret. For example, if p = 1, the model would have predicted every choice in the sample correctly. If p = 0, there would be no gain from changing any of the estimated parameters from zero.

The other measure for "goodness of fit" is described by Aldrich and Nelson (1984). They used the likelihood ratio statistic as follows:

$$C = -2 (\log LO - \log L1) \quad [5]$$

where,

L0 = value of log likelihood with parameters = 0

L1 = value of log likelihood with full model.

This follows a χ^2 distribution with k - 1 degrees of freedom (where K = number of coefficients contained to be zero) and tests Ho: all coefficients except the intercept = 0.

Therefore,

$$c = -2 (-119.05 - (-83.82)) = 70.46 \quad [6]$$

$\chi^2(15,0.05) = 25.00$ and Ho is rejected.

Model I can be considered a reasonable predictor of management behavior, but a number of the variables on the right-hand side had coefficients with incorrect signs and they were correlated with other variables.

The Stepwise procedure (Doan, 1989) was used to try to reduce the size of the model, but only two variables prevailed due to the low t - statistics (high standard errors) for the parameters in the model. Also, in another attempt to reduce the model size, a small model was estimated (2 variables) using the parameters that the author believed would most influence the decision to manage. The next most important variable was added and the model was estimated (3 variables) again and so on. This procedure did not statistically improve the model either. Both of these procedures were abandoned.

The next step was to identify highly correlated variables and, through an elimination process, Model II was estimated. The equation for Model II is as follows:

$$\log \frac{P(manage)}{1-P(manage)} = \quad [7]$$

- 3.1230 + 0.1781 AGE + 0.1409 DIST
(1.4560)* (0.1757) (0.1805)

- 0.3967 FARM + 0.5783 HAVESOLD - 0.0125 INSTRUCT
(0.5158) (0.5081) (0.9494)

+ 0.6773 HAVESOLD + 0.0053 INSTRUCT - 0.1911 OBTAIN
(0.5293) (0.1001) (0.3992)

- 0.0867 OBTAIN - 1.1946 SCHOOL12 + 0.0003 TOTACR
(0.3705) (0.4529)* (0.0004)

+ 0.2250 TOTINC - 0.3203 WOODINC
(0.0005) (0.2154)

The only variables that were significantly (5%) different from zero were the intercept and SCHOOL12.

likelihood ratio index:

$$p = 1 - \frac{(-89.60)}{(-119.05)} = 0.25 \quad [8]$$

likelihood ratio statistic:

$$c = -2(-119.05 - (-89.60)) = 58.9 \quad [9]$$

$\chi^2(9,0.05) = 16.92$ and Ho is rejected

Model II correctly classified 91.6% of the data used in model construction.

Table 2. Likelihood ratio (c) tests on different variable combinations testing Ho: variables = 0.

Ho	c	χ^2(K-1,.05)	Result
INSTRUCT=OBTAIN=0	0.08	3.84	Accept Ho
INSTRUCT=OBTAIN=DIST=FARM=0	2.18	7.81	Accept Ho
INSTRUCT=OBTAIN=FARM=0	0.60	5.99	Accept Ho
INSTRUCT=OBTAIN=FARM=DIST=AGE=0	3.46	9.49	Accept Ho
HAVESOLD=SCHOOL12=WOODINC=0	16.32	5.99	Reject Ho*
INSTRUCT=OBTAIN=DIST=0	1.36	5.99	Accept Ho
HAVESOLD=SCHOOL12=WOODINC=TOTING=0	48.78	7.81	Reject Ho*
HAVESOLD=SCHOOL12=WOODINC=TOTINC=TOTACR=0	49.42	9.49	Reject Ho*

The likelihood ratio statistic was then applied to Model II to test if there were variable combinations that equalled 0. The results of these tests are summarized in Table 2.

The null hypothesis was accepted five out of the eight tests (Table 2). Based on these results a third Model (Model III) was estimated dropping the variables INSTRUCT, OBTAIN, DIST, FARM and AGE. This model was chosen as the final model.

Results

Model III, the final model, consists of only five variables plus the intercept. The equation for the model is:

$$\log \frac{P(\text{manage})}{1-P(\text{manage})} = \quad [10]$$

$$\underset{(0.9339)^*}{-2.6306} + \underset{(0.4938)}{0.6901}\ \text{HAVESOLD}$$

$$\underset{(0.4305)^*}{-0.9961}\ \text{SCHOOL12} + \underset{(0.0004)}{0.0004}\ \text{TOTACR}$$

$$\underset{(0.1816)}{+0.1460}\ \text{TOTINC} \underset{(0.2764)}{-0.3121}\ \text{WOODINC}$$

The same two variables that were in Model II, intercept and SCHOOL12, are significantly different from zero (5%).

likelihood ratio index:

$$p = 1 - \frac{(-89.60)}{(-119.05)} = 0.25 \quad [11]$$

likelihood ratio statistic:

$$c = -2\,(-119.05 - (-91.32)) = 54.46^* \quad [12]$$

$\chi^2(4,.05) = 9.44$ and Ho is rejected

The final model correctly classified 91.7% of the data used in model construction. Model II contained only three variables that could be compared to the Clements and Jamnick (1989) model. These variables are shown in Table 3.

The WOODINC variable has an opposite sign compared to the Clements and Jamnick (1989) result. One would have thought that as more income was derived from the woodlot, the probability of that owner deciding to manage would increase. This does not appear to be the case in Carleton and Victoria counties. Also, a notable variable that was dropped from Model II was FARM. Even though a high percentage of owners in this area are farmers, and previous studies show that farmers are more inclined to decide to manage, this again does not appear to be the case for this area. It should be noted that only approximately 30 owners actually "managed" their woodlot from the survey data. This is a very low participation rate (7%) and may account for some of the differences exhibited between the results in this study and Clements and Jamnick (1989).

Table 3. Model III variables compared to the results from Clements and Jamnick (1989).

Variable	CVWPA		Clements and Jamnick	
	Coeff	Std. error	Coeff	Std. error
HAVESOLD	0.6901	0.4988	1.0934	0.0817
TOTALINC	0.1460	0.1816	0.1431	0.0348
WOODINC	-0.3121	0.2764	0.2362	0.0714

Conclusions

The final model in this study provides some interesting observations.

First, the owners in the CVWPA do not have the same profiles as the Clements and Jamnick's province-wide profile (1989) when it comes to the decision to manage. This indicates that province-wide results or policy should be implemented with caution. Policy makers should carefully analyze what type of programs should be implemented to encourage management. What makes intuitive sense or basing policy on results of previous, similar surveys elsewhere, may not provide the expected results. In other words, regional differences probably exist, and these differences possibly warrant regional policies.

Second, from Model III, the probability that an owner will decide to manage in the CVWPA area increases with the owner selling forest products, if total area owned increases and if his total income increases. The probability decreases if the education level is grade 12 or less and as wood income increases.

Finally, similar surveys and analyses should be carried out in other woodlot owner association or Marketing Board jurisdictions.

References

Aldrich, J.H., and Nelson, F.D. 1984. Linear Probability, logit, and probit Models. Sage University Paper Series on Quantitative applications in the Social Sciences, 07-045. Beverly Hills, New Delhi and London: Sage Pubns.

Clements, S.E., and Jamnick, M.S. 1989. Forest Management on Private Woodlots in New Brunswick. J. Environ. Manage. 29: 391-398.

Doan, T.A. 1989. Users Manual Rats Version 3.02. VAR Econometrics, Inc. Evanston, IL 60201.

Gramann, J.H., Marty, T., and Kurtz, W.B. 1985. A logistic analysis of the effects of beliefs and past experience on management plans for non-industrial private forests. J. Environ. Manage. 20: 347-356.

Pindyck, R.S., and Rubinfeld, D.L. 1991. Econometric Models and Economic Forecasts. Third Edition, McGraw-Hill, Inc. pp. 596.

Statistics Canada. 1990. Provincial Gross Domestic Product by Industry. Catalogue 15-203: Annual.