

University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

1989

Quantitative analysis of sold and unsold Forest Service timber offerings in Region 1

Michael J. Niccolucci
The University of Montana

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Niccolucci, Michael J., "Quantitative analysis of sold and unsold Forest Service timber offerings in Region 1" (1989). *Graduate Student Theses, Dissertations, & Professional Papers*. 3341.
<https://scholarworks.umt.edu/etd/3341>

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

COPYRIGHT ACT OF 1976

THIS IS AN UNPUBLISHED MANUSCRIPT IN WHICH COPYRIGHT
SUBSISTS. ANY FURTHER REPRINTING OF ITS CONTENTS MUST BE
APPROVED BY THE AUTHOR.

MANSFIELD LIBRARY
UNIVERSITY OF MONTANA
DATE: 1989

A Quantitative Analysis of Sold and Unsold
Forest Service Timber Offerings in Region 1

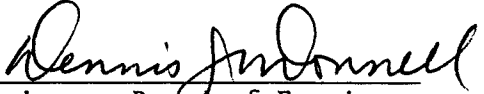
By

Michael J. Niccolucci

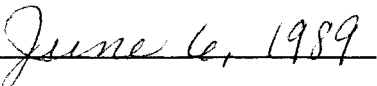
B.A., University of Montana, 1982

Presented in partial fulfillment of the requirements
for the degree of
Master of Arts
University of Montana
1989

Approved by


Chairman, Board of Examiners


Dean, Graduate School


Date

UMI Number: EP34496

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI EP34496

Published by ProQuest LLC (2012). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

7-17-89

Niccolucci, Michael, J., M.A., June 1989

Economics

A Quantitative Analysis of Sold and Unsold Forest Service Timber Offerings in the Northern Region (57 pp.)

Director: Dennis J. O'Donnell *DJO*

The United States Forest Service sells timber in the form of timber sales. In the 1980's the Forest Service has experienced an increase in the number of timber offerings that receive no bid, i.e. an unsold offering. During sample period, 1980 through 1985, certain national forests have experienced 25% of their offerings receiving no bid. The Northern Region in 1984 reached a high of approximately 24% of their volume offered receiving no bids. These unsold offerings, given all of the hours of planning needed in order to generate a timber sale that is ready for auction, are considered wasteful in terms of dollars and manpower. Modeling the occurrence of unsold timber offerings was the central issue of this work.

The characteristics of the timber sale, the timber sale contract, and the local and national economy were used to specify the classification models. The statistical models, Discriminant Analysis and Logistic Regression, were used to classify timber offerings into the categories of sold or unsold timber offerings. Also, given the large number of characteristics that describe a timber offering, Principal Components were developed to generate a smaller, parsimonious set of attributes.

The final models classified approximately 80% of the sales correctly. These models provide accurate information concerning the probability of a timber sale offering selling, given timber sale and economic characteristics of the offering. The models could prove to be an effective planning tool for land managers.

TABLE OF CONTENTS

ABSTRACT	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	v
ACKNOWLEDGEMENTS	vi
1. INTRODUCTION	1
1.1. PLANNING THE TIMBER SALE	1
1.2. OFFERING THE TIMBER SALE - FAILURE TO SELL	2
1.3. AUCTION THEORY - A BRIEF OVERVIEW OF THE LITERATURE	4
1.4. GENERAL MODEL SPECIFICATION	10
2. STATEMENTS OF HYPOTHESES	12
2.1. HYPOTHESES	13
2.1.1. SALE CHARACTERISTICS	13
2.1.2. ADMINISTRATIVE CHARACTERISTICS	16
2.1.3. ECONOMIC CHARACTERISTICS	17
3. METHODS	20
3.1. SAMPLING	20
3.2. MODEL SELECTION CRITERIA	21
3.3. DISCRIMINANT ANALYSIS	21
3.4. QUALITATIVE RESPONSE MODELS	25
3.4.1. LINEAR PROBABILITY MODEL	26
3.4.2. LOGIT VERSUS PROBIT	26
3.5. PRINCIPAL COMPONENTS	30
4. EMPIRICAL RESULTS	33
4.1. LINEAR DISCRIMINANT FUNCTION RESULTS	33
4.1.1. SALE CHARACTERISTICS	34
4.1.2. ADMINISTRATIVE CHARACTERISTICS	35
4.1.3. ECONOMIC CHARACTERISTICS	35
4.1.4. GOODNESS OF FIT RESULTS	36
4.2. LOGISTIC REGRESSION RESULTS	37
4.2.1. SALE CHARACTERISTICS	38
4.2.2. ADMINISTRATIVE CHARACTERISTICS	39
4.2.3. ECONOMIC CHARACTERISTICS	39
4.2.4. GOODNESS OF FIT RESULTS	39
4.3. PRINCIPAL COMPONENTS MODEL	40
4.3.1. LINEAR DISCRIMINANT FUNCTION	41
4.3.2. LOGISTIC REGRESSION	42

5. CONCLUSIONS	45
5.1. FURTHER RESEARCH	47
BIBLIOGRAPHY	50

LIST OF TABLES

TABLE 1-1:	REGION 1 TIMBER HARVEST	1
TABLE 1-2:	REGION 1 VOLUME SOLD AND VOLUME NOT SOLD	3
TABLE 4-1:	LINEAR DISCRIMINANT MODEL	34
TABLE 4-2:	DISCRIMINANT MODEL CLASSIFICATION RESULTS	36
TABLE 4-3:	DISCRIMINANT MODEL JACKKNIFE CLASSIFICATION RESULTS	37
TABLE 4-4:	LOGISTIC REGRESSION MODEL	38
TABLE 4-5:	LOGISTIC REGRESSION CLASSIFICATION RESULTS	40
TABLE 4-6:	DISCRIMINANT MODEL USING PRINCIPAL COMPONENTS	42
TABLE 4-7:	LOGISTIC REGRESSION MODEL USING PRINCIPAL COMPONENTS	43

ACKNOWLEDGEMENTS

There are many people that must be thanked. I first give thanks to my committee, Dr. Dennis O'Donnell, Dr. David Jackson and Dr. Thomas Power for being patient with a graduate student who works at his own pace.

Special gratitude goes to Dr. Ervin G. Schuster for not giving up on me as the months turned into years. Also, for being the individual who opened the doors within the National Forest System which enabled data collection to progress so smoothly.

Finally, to my lovely wife, Mary, who was always ready to listen and who gave her unending support.

CHAPTER 1
INTRODUCTION

The Western Montana and Northern Idaho area is the prime timber producing area of the Northern Region of the United States Forest Service. Within this area can be found a very diverse group of timber purchasers. They range from large national wood products firms like Champion International and Louisiana Pacific, to small private loggers (gypos). A major source of their timber supply, the input into their production process, is found on federal lands. There are major holdings of industrial and nonindustrial timber, but the United States Forest Service is still the major supplier (see Table 1.1 below).

TABLE 1-1:

REGION 1 TIMBER HARVEST (MMBF)

YEAR	TIMBER HARVEST FOREST SERVICE	FOREST SERVICE PERCENTAGE OF TOTAL	TOTAL HARVEST
1980	822.1	36.6	2247.8
1981	853.8	42.9	1988.2
1982	523.4	29.9	1751.6
1983	1061.0	42.1	2522.4
1984	907.7	35.9	2526.2
1985	983.4	41.8	2352.4

1.1. PLANNING THE TIMBER SALE

The Forest Service sells timber from federal lands in the form of timber auctions. The time span encompassing timber sale development typically ranges from approximately two to ten years. The first step in this process is for a field forester to walk through a particular area

and determine whether the area can support a timber sale; meaning that the trees in the area are mature and ready to harvest. The area is reviewed by foresters, engineers, and planning team members to determine how best to package the timber into a timber sale. The planning team makes decisions concerning total volume harvested and how it will be harvested. For example, in some areas, maintenance of certain visual quality objectives is desirable. If the sale is in one of these areas, adequate standing timber will be left on the land and the shape of the cutting units will be designed to mesh with the natural landscape. Decisions are made concerning whether road building is necessary, and if so, what road standards should be used. If the soil is erosive, both road building and the yarding method (the means used to get the logs from the stump to the landing site) may be altered to reduce erosion. There are also considerations concerning wildlife. "Is the area summer or winter range for elk?" "Is it elk calving habitat or is there any evidence of rare and endangered species?" If any of these are true, logging plans may be further modified.

Timber sales can be used to produce positive externalities. The timber sale, for instance, may be used to benefit wildlife, recreation and other nontimber outputs (Lyon et al, 1985).

1.2. OFFERING THE TIMBER SALE - FAILURE TO SELL

Once the timber sale planning is finalized, the sale is offered to the public. The outcome desired by the Forest Service is for the sale to sell and for the return in stumpage sold to cover their costs of planning and administration. However, sometimes the result, especially

in the 1980's, is that the timber sale offering will result in a no bid sale, i.e. an unsold offering. This result, given all of the hours of planning needed to generate a timber sale that is ready for auction, is wasteful in terms of dollars and manpower. It is not, as some Forest Service officials claim, something that can be put on a shelf and stored. Many of these timber sales will need to be revised in order to arrive at an offering that has a high probability of selling. Of course these revisions require more money, and there is no guarantee that the revisions produce a sale that is more appealing to a potential buyer. Table 1.2 presents the problem of unsold offerings in terms of volume offered, volume sold, and volume unsold over the period of the study, 1980 - 1985.

TABLE 1.2:

REGION 1 VOLUME SOLD AND VOLUME NOT SOLD (MMBF)

YEAR	VOLUME OFFERED	VOLUME SOLD	VOLUME UNSOLD
1980	1287.0	1096.2	190.8
1981	1223.4	969.7	253.7
1982	1179.5	951.0	228.5
1983	1223.6	1048.0	175.6
1984	1169.5	878.5	291.0
1985	1145.7	914.4	231.3

The failure of sales to successfully sell raises the obvious question: "Why do these timber sale offerings go unsold?" Can the problems be traced to the characteristics of the timber sale or to problems in the economics of the wood products industry or to a combination of both? Modeling this phenomena will be the central issue of this thesis. Hopefully as a result of this analysis, a policy tool

will result that allows planners to quantify the likely consequences of their timber sale planning decisions.

1.3. AUCTION THEORY - A BRIEF REVIEW OF THE LITERATURE

An auction process has long been considered an efficient means of allocating resources. The auction market not only allocates resources, but also generates information that is of value for both buyer and seller. The English auction is used by the USDA Forest Service, R-1, as a means of selling timber sales. The English auction starts at a low price level, and then gradually rises until only one willing bidder remains (Milgrom and Weber, 1982). In Forest Service timber sale auctions, the starting point of the auction is the advertised rate. Table 1.2 indicates that buyers of Forest Service timber sales are generating important information concerning timber sales by not bidding on approximately 20% of the volume. The following will review three theoretical auction frameworks set forth in the economic literature, and then, review several empirical works that focus on timber sale auctions.

The first theoretical auction framework is the Independent Private Values (IPV) model. The IPV model suggests that each bidder knows his own valuation with certainty, but not those of his competitors (this is the private values assumption). "When firms are bidding for inputs, differences in the good's subjective values may be due to differences in location, product mix, and resources" (Brannman, Klein and Weiss, 1987, p. 24). These bids are considered independent and drawn from a known value distribution which is known by all auction participants. "Bidders are assumed to behave competitively; therefore, the auction is treated

as a noncooperative game among the bidders" (Milgrom and Weber, 1982, p. 1090). The outcome using the IPV model, under an English auction is Pareto optimal; that is, the winner is the bidder who values the object most highly.

The second theoretical auction framework set forth is the Common Value (CV) model. "The CV model assumes that the object offered for sale has some "true" value which is common but unknown to all bidders" (Branman, Klein and Weiss, 1987, p. 25). The CV model allows for statistical dependence among bidders' value estimates, however offers no role for differences in individual tastes (Branman, Klein and Weiss, 1987). The CV model is susceptible to the winner's curse. The winner's curse arises when more than one bidder bids more than his estimate of the good's value. The winner's curse can be avoided if more cautious bids are submitted. More caution in the auction process means lower bids.

The Milgrom and Weber framework (MW) is not as restrictive as the IPV and CV models. MW argues that the IPV model does not allow for uncertainty concerning the goods true value to enter the process. It also considers the bids to be interdependent, not independent. Under the CV model, the restriction does not allow differences in tastes. Given these criticisms, the MW model allows elements of both IPV and CV to enter the process. However, as competition increases it is expected that the winning bids in the MW model should lie between the IPV and pure CV predictions.

Haynes (1980) analyzed the competition for Forest Service timber sales. The study covered three regions, Pacific Southwest, Pacific Northwest and the Northern Region. Throughout the research the assumptions of competitive markets was used. "A competitive market model assumes that neither the buyer nor the seller can influence price through the sizes of their purchases or sales" (Haynes, 1980, p. 8). With the assumption of a competitive market as a basis, sales are delineated into two groups, competitive and noncompetitive. The delineation is based upon a price variable, overbid. Overbid is equal to the bid price minus road costs and appraised stumpage (Haynes, 1980). Haynes argues that overbid provides an absolute and consistent measure of competition in comparison with a bid-appraisal ratio used in other research (see Mead, 1981). Using overbid as a measure of competition, a sale is defined to be noncompetitive if the overbid was less than one-half of 1 percent of the average overbid for the appraisal zone in which the sale is located; and competitive if the overbid is greater than one-half of one percent of the average overbid for the appraisal zone in which the sale is located. The assumption is that noncompetitive sales are somewhat different in terms of physical and administrative characteristics. These physical and administrative characteristics, along with appraisal information, were used to measure the quality of the sale. When faced with a sale of questionable quality, bidders will not increase their bid above the minimum acceptable bid, advertised rate. Bidders, by placing a low bid, are expressing their belief that the sale is of questionable quality and

will result in low or no profitability for the firm. Discriminant analysis, using physical, administrative, and appraisal information as independent variables, was used as the classification tool. The discriminant models provided an objective means of classifying timber sales into the classes of competitive and noncompetitive. The results from the three Regions were quite similar. In general, timber sales with lower volumes per acre and lower grade species, but higher road and logging costs, are more likely to be noncompetitive sales. These characteristics identified a sale of low quality, and therefore, of low or no profitability.

Johnson (1979) analyzed sold timber sales during July 1, 1973 to July 1, 1975 in the Northern Region. His "major objective was to show that there are certain circumstances whereby oral and sealed bidding can be expected to yield the same price and other circumstances where sealed bidding is expected to yield a higher price" (Johnson, 1979, p. 316). Johnson delineates the two cases by distinguishing between two bidding cases, homogenous and nonhomogenous. The homogeneous bidding case is derived in the following manner. Each bidder evaluates the quality of the timber on the sale site. Each bidder determines the products that can be produced from the timber on the sale site. Given that the final products can be determined, an end product value (P_i) follows. Each bidder has a known cost structure, C_i , for manufacturing and logging. However, the true logging and manufacturing costs are not known. "Given uncertainty and errors in estimation, both P_i and C_i can vary from bidder to bidder" (Johnson, 1979, p. 317). "The maximum value, V_i ,

that any bidder (i) is willing to pay for the timber is, $V_i = P_i - C_i$ " (Johnson, 1979, p. 317). The homogeneous bidding case is where each bidder has an equal chance of drawing the highest V_i and winning the auction (Johnson, 1979). The nonhomogenous case is when a bidder or group of bidders will be known to have a lower cost structure for a particular type of timber sale. The criteria for identifying sales where bidders don't have an equal chance of winning the timber sale auction is based upon the road building requirements.

"The most important existing constraints on sale size are those imposed by the necessity of road development funding through timber purchases. The high per mile costs ... mean that a sufficient volume of timber must be included in a sale to enable the purchaser to recover his road building costs through stumpage payment credits. This is a particular problem in areas with high road costs and low values per unit of timber volume. As a result, the smaller purchasers are often incapable of handling such sales because of the working capital requirements and specialized equipment needed" (Johnson, 1979, 320).

Sales with less than \$20,000 of road building are considered to be in the homogeneous class; and sales with more than \$20,000 of road building are in the nonhomogeneous class. Sales with smaller road developments have the likelihood of receiving more bids because a larger number of potential purchasers exist that can harvest the sale profitably. More firms (both small and large) possess the necessary capital equipment needed to harvest the sale profitably. However, sales with large road building requirements have a smaller number of potential purchasers that can profitably bid on the sale because of the ownership of the necessary capital requirements and specialized equipment needed.

Merzenich (1981) in his developmental transaction evidence appraisal work in the Northern Region found that when skyline yarding was

required, a lower number of bidders participated in the timber sale auction. Skyline yarding systems are specialized systems, which are not owned by all purchasers, especially small purchasers. Like specialized road building equipment, required skyline yarding systems reduce the number of potential buyers, and translate into higher costs. If the timber sale is composed of lower valued species and specialized yarding systems, the profitability of the sale is likely to be small. However, as in the case of road building equipment, purchasers who own specialized yarding systems will have the advantage of a lower cost structure for sales of this type.

Jackson and McQuillan (1979), using a sample of timber sales from three districts on the Lolo National Forest, developed a hedonic price function. The model predicted stumpage price based upon tree diameter, stand composition, logging method, reproduction system and market conditions. They hypothesized that net stumpage value for trees smaller than a given diameter, woodland processing and delivery costs exceed the log values delivered at a mill (Jackson and McQuillan, 1979). For trees with larger diameters, a positive value exists, "because per-unit felling, bucking, skidding and loading costs decrease with increasing diameter while realized value at a mill increases with size due to reduced handling costs and higher product values" (Jackson and McQuillan, 1979, p. 621). As in Haynes' research, Jackson and McQuillan use sale characteristics as proxies of sale quality. In general, the model indicates that sales which have on the average smaller diameter timber (lower value), lower volumes per acre (less value per acre),

skyline yarding (higher yarding costs), and are sold during depressed markets (depressed lumber prices), will have a lower sale value. The empirical model was highly successful at explaining the variation found in bid price, $R^2 = .81$.

The auction frameworks set forth describe the theoretical foundations that can be expected under certain assumptions. Timber sales are a resource that possess significant uncertainty. The uncertainty stems from a lack of knowledge with respect to the quality of the timber to be harvested, the exact cost of harvesting the sale, and the volatility of end product prices. The empirical auction research centered on using these quality indicators to model various aspects of federal timber sales.

1.4. GENERAL MODEL SPECIFICATION

The model specification was based on the knowledge generated from the literature review. None of the literature cited specifically dealt with unsold sales. However, unsold sales can be considered a special category of Haynes' noncompetitive sales category.

The empirical research reviewed above used quality measures in their model specifications. In general, the quality measures reflect the revenues that are expected to be generated from the stumpage versus the costs of harvesting the stumpage. Given the objective of this research, understanding and modeling of unsold versus sold sales, the general model specification was the following:

Sold and Unsold Sales = f(1. Sale characteristics
 2. Administrative characteristics
 3. Economic characteristics).

The specific variables that were used to represent the above model specification are presented in the following chapter.

CHAPTER 2

STATEMENTS OF HYPOTHESES

This chapter will deal with the identification of the variables that could be used in the models and statements of the hypotheses about the expected signs on these variables.

The variables that were used to predict sold and unsold timber sales can be grouped into three categories: sale characteristics, administrative characteristics and economic characteristics. Each of these categories represent specific attributes of the timber sale. Sale characteristics represent physical features of the timber sale; administrative characteristics represent the contractual agreements of the timber sale; and the economic characteristics represent the economic situation at the time of the auction. These attributes are proxies for the items that increase (decrease) revenues or increase (decrease) costs. For example, as stump to mill costs increase, holding revenues from the sale constant, it is more likely that the prospective purchaser of the sale will not purchase the sale because of the increased costs (decreased revenue) that are likely to result. Also, when a sale characteristic like percent volume skyline yarded is used in the model, it is a proxy for increased costs. Skyline yarding is one of the most expensive yarding methods available; when it is required this, translates into logging in difficult terrain.

One tailed tests of significance were used since I hypothesized the sign as well as the non-zero significance of each variable. One must remember that these hypotheses are made under the assumptions that all other variables in the model are held constant. However, in a model that is a linear combination of a set of variables, it is important to remember that we are making a conditional probability statement. The hypothesis is conditional on the other variables found in the equation. Even more caution should be used when certain variables are correlated. These hypotheses are derived from my knowledge of economics and forestry, and are also backed by what is observed in the empirical forest economics literature.

2.1. HYPOTHESES

2.1.1. SALE CHARACTERISTICS

1) Stump to Mill Costs:

$$H_0: B = 0$$

$$H_a: B < 0$$

Stump to mill costs are defined to be Felling and Bucking costs, Skidding and Loading costs, Haul costs, Slash Disposal costs, Specified Road costs and the Advertised Rate. These activities are the necessary steps the purchaser needs to complete in order to make the standing trees into log deck inventory. The purchaser must buy the rights to the trees with the minimum price as the advertised rate, build the necessary roads (Specified Road costs), fell the trees (Felling and Bucking), skid the logs to a landing (Skidding and Loading), dispose of the slash (Slash Disposal) and haul the logs to the mill (Haul costs). These are Forest Service appraisal estimates that are found on the appraisal form 2400-17. These appraisal estimates reflect the average costs experienced by a sample of purchasers in the region. It is important to remember that these cost items are based on average costs, not marginal costs and therefore, may not be accurate estimates of the winning bidders costs of logging.

2) Total Harvested Acres

----- = Dense Pack:

Total Sale Acres

$$H_0: B = 0$$

$$H_a: B > 0$$

Dense Pack^a is measuring the amount of movement of people and equipment that is needed to harvest the given sale volume. As this ratio approaches one, I hypothesize that costs of harvesting a certain volume decreases. With stump to mill costs decreasing, the sale should become more appealing.

3) Average diameter at breast height:

$$H_0: B = 0$$

$$H_a: B > 0$$

As the average diameter increases, the value of the logs increase, this should decrease the chances of a no-bid sale.

4) Average Logs/Thousand Board Feet:

$$H_0: B = 0$$

$$H_a: B < 0$$

This variable indicates the number of logs that need to be moved in order to harvest a thousand board feet of timber. The more pieces that need to be handled, the higher the stump to mill costs. Logs/thousand board feet is highly correlated with average diameter at breast height. I will only use one of these variables in the equation in order to eliminate any problem with collinearity. As logs/thousand board feet increases, this will reduce the chances of the sale selling.

5) Volume/Acre Harvested:

$$H_0: B = 0$$

$$H_a: B > 0$$

As volume per acre increases this will decrease felling and bucking costs and skidding and loading costs per acre harvested. The logger will have to move less in a given area in order to harvest more volume. Thus, the more volume/acre the better the chances of the sale selling.

6) Total Haul Miles to the first Appraisal Point

----- = Haul Ratio

Total Haul Miles to the next Appraisal Point

$$H_0: B = 0$$

$$H_a: B > 0$$

This ratio represents how isolated the sale is from various milling centers. The numerator is the haul distance to the Forest Service appraisal point. The denominator is the haul to the next nearest appraisal point. If the ratio is small, this indicates the sale is isolated from potential milling areas. If the ratio is close to one, the sale has at least two potential milling areas as outlets.

7) Percent of the volume tractor yarded:

$$H_0: B = 0$$

$$H_a: B > 0$$

Tractor yarding is the cheapest method available to yard logs. If a sale has a higher percentage of the volume tractor yarded this should decrease skidding and loading costs, which will then make the sale more appealing. Having sales that are 100% tractor yarded also increases the number of potential bidders because all loggers have tractors, but not all have cable systems.

8) Percent of the volume cable yarded:

$$H_0: B = 0$$

$$H_a: B < 0$$

Cable yarding is more costly than tractor yarding and will increase the cost of yarding. And as stated above, fewer loggers have cable systems reducing the number of potential bidders.

9) Percent of the sale volume clearcut, seedtree cut or right of way:

$$H_0: B = 0$$

$$H_a: B > 0$$

These silvicultural methods allow the logger to harvest the most volume per acre, thus, reducing felling costs per acre. In Jackson's study (1987) comparing Forest Service and Department of State Lands bid values, the estimated coefficient was negative. This is the opposite sign that I hypothesized above. The negative sign is not surprising if the quality of the trees is considered. If lower valued species, dead, or diseased trees are being harvested then the negative sign is justified.

10) Miles of new specified roads:

$$H_0: B = 0$$

$$H_a: B < 0$$

As the miles of new road construction increase, this will decrease the probability of the sale selling. The reason is, it delays the logging operation, sometimes for as long as two years. This allows the problem of business cycle risk to enter the process. Another view is presented by Johnson (1979). Johnson classifies timber buyers into 2 groups, homogeneous and nonhomogeneous buyers. The membership in these groups depends upon the degree to which firms are similar or different in regards to ownership of certain capital assets (Johnson, 1979). A timber buyer who owns the necessary capital equipment in order to build roads in a timber sale will have a distinct bidding advantage. This will then cause the above hypothesis to be positive.

11) Miles of reconstructed specified roads:

$$H_0: B = 0$$

$$H_a: B < 0$$

The comments for miles of new specified roads apply here.

2.1.2. ADMINISTRATIVE CHARACTERISTICS

1) Contributed funds:

$$H_0: B = 0$$

$$H_a: B > 0$$

If funds are contributed to the sale to cover the cost of the roads, this should increase the likelihood of the sale selling. The Forest Service has eliminated any risk associated with the road building for the potential buyer. This will then make the sale more attractive.

2) Small Business Administration Setaside Sale:

$$H_0: B = 0$$

$$H_a: B \neq 0$$

Two views can be considered. The first is that it reduces the number of potential bidders; this will make it more likely to not sell. Second, this increases competition for the sale, by giving the small purchaser an advantage. An SBA setaside sale is not randomly chosen. Forest officials decide if a sale will be advertised as a SBA based on whether SBA size firms (500 employees or less) have been successful in purchasing a certain percentage of the sales in a particular forest.

3) Escalation Price Provision:

$$H_0: B = 0$$

$$H_a: B > 0$$

In a sale that uses an escalated price clause, the stumpage price is linked to the WPPA price index. If the price falls, the purchaser receives a full discount on stumpage; whereas if prices rise, the purchaser pays an added price commensurate with only 50% of the increase. Escalated sales decrease the risk to the purchaser, and this should make the sale more appealing. Escalation can only apply to sales that are not appraised as deficit sales.

4) Contract length in months:

$$H_0: B = 0$$

$$H_a: B \neq 0$$

A specific sign will not be stated for this variable because of the changes that occurred in the wood products industry during the early 1980's. At one time, a long contract period could be translated into greater flexibility to the purchaser. But today, long contract periods could be viewed as greater risk because of the problems in the wood products market. Also, contract length is correlated with total volume, miles of road constructed and total acres harvested. This correlation is created by the regulations found in the Forest Service Manual.

2.1.3. ECONOMIC CHARACTERISTICS

1) Selling Value (Lumber Tally):

$$H_0: B = 0$$

$$H_a: B > 0$$

Selling Value L.T. (SPLT) is the value of the product that can be produced from the given log species input. This is not a perfect estimator of the value of the lumber that can be produced from the logs, as the Forest Service assumes that the producer will maximize the volume of lumber that can be produced from the logs. The producer may maximize profits by maximizing the output of specific types of products, which may not lead to the maximum lumber output. Given this slight drawback, SPLT is still a good measure of lumber price.

2) Monthly Housing Starts:

$$H_0: B = 0$$

$$H_a: B > 0$$

Since a majority of the output from Region 1 mills is housing construction products this economic measure will give an indication of the strength of the demand for raw materials.

3) Home Mortgage Interest Rates:

$$H_0: B = 0$$

$$H_a: B < 0$$

Home mortgage interest rates is the cost of financing a new home to the consumer. High rates tend to reduce the demand for new housing, thus depressing the demand for stumpage.

4) Framing lumber composite prices:

$$H_0: B = 0$$

$$H_a: B > 0$$

If the framing lumber composite price is rising this an indication of good lumber markets. If lumber markets are perceived to be rising, and this rise is expected to continue, this will spur timber buyers to buy timber sales.

5) Uncut Volume Under Contract:

$$H_0: B = 0$$

$$H_a: B < 0$$

Uncut volume under contract is the amount of volume purchasers have under contract with national forests. It can be considered their savings account. If the purchaser possesses a large amount of uncut volume under contract, he will be more selective when purchasing new sales. The purchaser knows mill needs will be met by his uncut volume under contract and other sources, thus, he will forgo timber sales that are of questionable value.

6) Lumber orders by mills in the Intermountain region:

$$H_0: B = 0$$

$$H_a: B > 0$$

If mills are filling more orders, this indicates an increase in demand. With an upswing in this measure, mills will be more willing to add logs to their inventory.

7) Mill Production in the Intermountain region:

$$H_0: B = 0$$

$$H_a: B > 0$$

Mill production indicates a positive change in the lumber market (or a company increasing a depleted lumber inventory). This will lead to a hypothesized positive sign on this variable.

8) Mill 1:

$$H_0: B = 0$$

$$H_a: B > 0$$

This variable indicates if the appraisal point chosen by the Forest Service is competitive. A competitive appraisal point is one with more than one competing mill. If the appraisal point meets the above criteria then it is coded as a one, otherwise it is coded as a zero.

9) Canadian/United States dollar exchange rate:

$$H_0: B = 0$$

$$H_a: B < 0$$

The exchange rate is the price of one currency in terms of another currency, and in the long run, is determined by the general price level within the two economies (McCarl and Haynes, 1985). There has been much said concerning the effect Canadian lumber has had on U.S. markets. Most of the discussion has centered on the "subsidy" the Canadian timber purchasers receive when purchasing stumpage. Research has been done examining the issue from the aspect of exchange rates (McCarl and Haynes, 1985). During the early 1980's the U.S. dollar has been very strong, especially compared to the Canadian dollar. Exchange rates have the following two effects: 1. An increasing exchange rate encourages imports into the country and acts as an implicit import subsidy for foreign producers; 2. An increasing exchange rate discourages exports from the country and acts as an explicit tax on domestic producers (McCarl and Haynes, 1985). As this imbalance persists, Canadian lumber will continue to supply approximate 30% of the lumber sold in the United States, and thus, this will have a negative effect on the outcome of timber sale auctions.

Most of these variables have been used in other work dealing with timber sale economics (Buongiorno and Young, 1984; Connaughton, 1981; Jackson and McQuillan, 1979; Matthews, 1942; Merzenich, 1981;

Johnson-True, 1985). A much smaller number of variables were used in formulating the models. There are several reasons why all of the variables can't be used. First, the objective was to build a prediction model that is parsimonious. This means that the model should not be redundant in the variables that are present. Second, some of the variables stated above explain the same variation. This means these variables can be substituted for one another, and the final model will explain approximately the same amount of the variation. Third, if all of the variables were used in the modeling process, a degrees of freedom problem would arise.

The timber sale data can be found in the timber sale folders located at the Forest Supervisor's office and on computer file at the Regional office. The economic data can be found in published documents.

CHAPTER 3

METHODS

The motivating factor in this research was the development of a model (policy tool) that could be used to predict the outcome of a timber sale auction. It is important to have the ability to predict outcomes. However, it is also important to understand the phenomena that is being studied. The models developed in this research will allow the prediction of a timber sale auction, and also, will allow one to draw conclusions as to why unsold sales occur. Two statistical models were developed, Logistic Regression and Discriminant Analysis. In general, the logistic regression model predicts the probability of an outcome occurring (probability of a sale selling), and discriminant analysis classifies objects (timber sales) into groups (sold or unsold).

3.1. SAMPLING

A random sample of USDA Forest Service timber sales comprised the data for this study. It was viewed to be statistically beneficial to have a large sample of sold and unsold timber sales; therefore, a more intense random sampling was used for the unsold sales given that fewer of these sales occurred. The final sample consisted of 204 sold timber sales and 145 unsold timber sales from the seven westside forests (Bitterroot, Lolo, Flathead, Kootenai, Idaho Panhandle, Clearwater and Nezperce) in Region One during the years 1980 through 1985.

3.2. MODEL SELECTION CRITERIA

During model development, a model selection criteria must be developed which allows the modeler to objectively choose between models. The following items were used to decide which model was considered best:

1. How well does the model predict (goodness of fit) and;
2. What type of information does it provide to the user.

3.3. DISCRIMINANT ANALYSIS

In this analysis we have two populations, call them π_1 and π_2 . These two populations represent sold and unsold sales. The basic goal of Discriminant Analysis is to produce a linear function that will optimally separate these two populations. Each population has its own expected value $E(\mathbf{x}) = \mu_1$ and $E(\mathbf{x}) = \mu_2$, but, they share the same variance-covariance Σ matrix, (This is actually a pooled variance-covariance Σ matrix). This assumption of equal variance-covariance matrix may empirically be a heroic assumption. Research has shown that if the covariance matrices are not equal for the two populations then the linear discriminant function is not robust (Press and Wilson, 1978). If the assumption of equal variance-covariance is violated, then it is appropriate to replace the pooled variance-covariance matrix with the individual variance-covariance matrices. By making this substitution the method is no longer linear but is a quadratic discriminant function (Johnson and Wichern, 1982).

Like most of the multivariate techniques, the discriminant function is a linear combination of p variables. A discriminant score is produced from this linear combination as follows:

$$Y = \mathbf{l}'\mathbf{x} = l_1x_1 + l_2x_2 + \dots + l_px_p.$$

Each population has its own mean

$$\mu_{1Y} = E[Y|\pi_1] = \mathbf{l}'\boldsymbol{\mu}_1 \text{ and } \mu_{2Y} = E[Y|\pi_2] = \mathbf{l}'\boldsymbol{\mu}_2$$

The variance of Y is:

$$\text{Var}(Y) = \mathbf{l}'\boldsymbol{\Sigma}\mathbf{l} = \sigma_Y^2$$

The l 's, the coefficients of the linear combination, are chosen to maximize

$$(\mu_{1Y} - \mu_{2Y})/\sigma_Y^2.$$

The numerator can be thought of as the between population variance and the denominator, variation within the population.

$$\mathbf{l}' = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\boldsymbol{\Sigma}^{-1}$$

$$Y = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\boldsymbol{\Sigma}^{-1}\mathbf{x}$$

The above equation is known as Fisher's Linear Discriminant Function.

There are two major assumptions that are the foundation of Discriminant Analysis. They are:

1. the \mathbf{X} matrix is distributed as a Multivariate Normal;
2. and each population has equal variance-covariance matrices.

If all variables in the analysis are continuous and if the sample size is large, i.e. greater than 50, then it is safe to make the assumption that the distribution is multivariate normal.

In this research as in most economic research, there are several variables that are measured in discrete terms, zero or one. When this occurs the assumptions of multivariate normality and equal variance-covariance matrices are violated (Press and Wilson, 1978; Eisenbeis, 1977). For example, assume we have two variables. If we plot these variables, the plot should resemble an ellipse if the bivariate normal holds. If the linear discriminant function is to be successful, these two populations should be separable with a linear function. But if assumption 1 and 2 above do not hold, then it will require a nonlinear function to perform the separation (quadratic discrimination). Also, if the covariance matrices are not equal it is not appropriate to pool them. It is this pooled covariance matrix that is the central building block to the Fisher linear discriminant function. Since access to computer software that would have allowed the estimation of a quadratic discriminant function was lacking, the use of dichotomous variables was held to a minimum.

The above statements indicate that the linear discriminant function is based on very strict assumptions which are seldom observed. Research has investigated the robustness of the linear discriminant function when the above assumptions are violated (Press and Wilson, 1978; Eisenbeis, 1977). Robustness refers to the deterioration in error rates caused by using a classification procedure with data that do not conform to the assumption on which the procedure was based (Johnson and Wichern, 1982).

In discriminant analysis it is not possible to determine the significance of individual variables as in classical linear regression (Eisenbeis, 1977). There are a number of methods which attempt to determine the relative importance of individual variables (Eisenbeis, 1977). A few examples of these are the univariate F-statistic and the stepwise forward inclusion method. The univariate F-statistic indicates which means are significantly different. The problem is that the variables are treated independently. The F-statistic may indicate that a variable is insignificant. When combined with another variable this variable may be quite important. Collinearity may help the discriminating power of the function up to the point where it is no longer possible to invert the dispersion matrices (Eisenbeis, 1977). The stepwise procedures are based upon the conditional methods which take into account correlations among the variables (Eisenbeis, 1977). The method used was a stepwise inclusion method based upon Wilk's Lambda. Wilk's Lambda is a statistic used in testing the hypotheses of equal group means. All of the methods used for testing significance assume equal dispersion matrices.

Once the models were estimated, the question of which variables (characteristics) were most important in determining sold and unsold timber sales arises. This question was answered by using the standardized canonical discriminant function coefficients. The standardized canonical discriminant function coefficient represents the relative contribution of its associated variable to that function, when the sign is ignored. The

sign indicates if the variable is making a positive or negative contribution. The interpretation of these coefficients is comparable to the interpretation of beta weights from regression analysis.

Once the model was estimated, its predictive power was determined. An estimate of the model's error rate is a count of the number of correctly classified outcomes compared to the total number of cases. This ratio indicated the accuracy of the model. However, this measure of accuracy is considered to be upwardly biased (Lachenbruch, 1975). A model should be able to predict the data from which it was built fairly accurately. The jackknife procedure builds the model on $n-1$ observations and then predicts the observation that was left out. This is done for each observation and an error rate is calculated. The jackknife error rate estimate is considered a better estimate. Both error estimates were used in this research.

Given that the assumptions of multivariate normality and equal dispersion matrices were not met in this research, a second statistical tool was developed that has been found to be robust under the above violations (Press and Wilson, 1978).

3.4. QUALITATIVE RESPONSE MODELS

There are basically three common forms of probability function used in application (Ameniya, 1981).

Linear Probability (LP) Model: $F(w) = w$

Probit Model: $F(w) = \Phi(w) = \int_{-\infty}^w \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$

Logit Model: $F(w) = L(w) = e^w / (1 + e^w)$

3.4.1. LINEAR PROBABILITY MODEL

The LP model is nothing more than a linear regression model that can be estimated using either Ordinary Least Squares (OLS) or Weighted Least Squares (WLS). The most obvious defect with this model is that it does not constrain its predictions to the interval zero-one, as a probability should. One method that has been used to correct this problem is to truncate the prediction, that is, any prediction greater than one is considered to be a one and any prediction less than zero is considered to be zero. Also, when this function is estimated using OLS, the variance of the error is heteroscedastic, $\text{var}(u_i) = \sigma^2$ (Koutsoyiannis, 1977). This problem can be remedied by estimating the function using WLS. This approach still has the shortcoming of predicting outside the interval of zero-one. However, it is still considered a good starting point for the model building process (Ameniya, 1981). Since the coefficient estimates still are statistically unbiased, the objective is strictly prediction, and there is not access to software that allows estimation of a probit or logit model, the linear probability model is a viable option. Since heteroscedasticity affects the standard error, tests of significance are not appropriate.

3.4.2. LOGIT VERSUS PROBIT

Given the above problems, a technique that is based on a cumulative probability distribution function would be more appropriate. There are many types of cumulative probability distributions. The two most commonly used are the cumulative normal probability function (Probit) and the

cumulative logistic probability function (Logit). The Probit model is based on the standard normal probability density function, and the Logit model is based upon the logistic probability density function (Ameniya, 1981). Both of these distributions are symmetric around zero and have variances equal to 1 and $\pi^2/3$, respectively (Ameniya, 1981). The literature has shown that these two distributions are very similar, in fact, it is difficult to distinguish them statistically unless one has an extremely large number of observations (Ameniya, 1981). Given this and the fact that the Probit model is much more computer intensive when considering model estimation, this research used the Logit model.

Much of the earlier work done using the logistic regression model consists of data that can be grouped. For example, a model contains a variable that measures three dosage levels of insecticide. The data can be grouped into three categories, each with a certain sample size. When the sample size is large in each "cell" the estimation process is relatively straightforward. In economic research, many of the independent variables are continuous. This translates into one observation per cell; or in other words, only one choice is associated with each set of independent variables. It is possible to aggregate the continuous variable into categories. For example, if income is an independent variable in the model, it is possible to aggregate the income data into categories. However, aggregation leads to a loss of information. Fortunately a maximum-likelihood-estimation procedure can be applied to the model (Ameniya, 1981). The maximum-likelihood-estimation procedure

does not require that the data be grouped (aggregated). This allows each individual observation within the sample to have a distinct probability with which it is associated (Pindyck and Rubinfeld, 1981). Briefly, the maximum-likelihood method chooses among all possible estimates of the parameters those values which make the probability of obtaining the observed sample as large as possible (Koutsoyiannis, 1977). These maximum-likelihood-estimates have several desirable properties (Pindyck and Rubinfeld, 1981). All parameters are consistent and efficient asymptotically. All parameter estimators are known to be normal, so that the t-test can be applied in order to test for the statistical significance of coefficients. The discriminant function does not allow us to do this.

Another advantage of the logistic regression approach is that the coefficients can be interpreted. It is important to note that the coefficients of the logistic regression cannot be directly interpreted as you would interpret standard regression coefficients. The interpretation of the standard regression coefficient is, *ceteris paribus*, a one unit increase in the exogeneous variable will lead to a change in the endogeneous variable equal to the amount of the estimated coefficient, B_i . The B coefficient from the logit model indicates the increase in the log of the odds that an event (sold offering) will occur given a unit change in the independent variable X_i (Pindyck and Rubinfeld, 1981). It can be shown that a change in P_i , as a result of a change in X_i , is a function of both B_i and P_i (Pindyck and Rubinfeld, 1981).

$$P_i = B_i * [P_i * (1 - P_i)] .$$

This also implies that changes in independent variables will have their greatest impact on the probability of choosing a given option at the midpoint of the distribution (Pindyck and Rubinfeld, 1981). This can also be understood when viewing the cumulative logistic curve. At the low slopes near the endpoints of the distribution large changes in X are necessary to bring about small changes in probability. At the center of the distribution (probability equals .50) the slope of the cumulative logistic curve is the greatest; a change in the X will yield the greatest change in the probability.

To measure how well the logistic regression fits is not as straightforward as in classical linear regression. The R-squared from classical linear regression indicates how much of the variation found in the dependent variable is explained. In this case the dependent variable is a continuous random variable. Given that the dependent variable in my analysis was a zero or a one, other methods were used to measure goodness of fit.

A goodness of fit method that is used extensively in the field of meteorology is the Quadratic score (Winkler and Murphey, 1980). The following is the quadratic scoring rule for evaluating a model's prediction:

$$\frac{\sum_i [1 - 2 * (\hat{P} - E)^2]}{\text{number of timber sales}} = \text{Average score}$$

where i runs over all sales found in the sample, where E is a random variable with a value equal to one if an offering sells and equal to zero if the offering does not sell. If \hat{P} , the predicted probability from the logistic regression model, is near 1 and $E = 1$, the prediction is good. Likewise, if \hat{P} is near zero and $E = 0$, the prediction is good. The range of possible values for this quadratic score is -1 to $+1$, with values being close to 1 being good scores and values near -1 being poor scores.

Another method used to test goodness-of-fit is identical to the method used for discriminant analysis. The number of correct predictions divided by the number of total cases will provide an error rate estimate. The problem with this measure is that cases are equally penalized. For example, predictions of $.01$ and $.49$, assuming the decision rule is $\geq .50$ equals sold and $< .50$ equals unsold, are equally predicted to be unsold, when actually the prediction of $.01$ is truly a correct prediction while the $.49$ could be considered an indecision (Ameniya, 1981). Like the discriminant function, the logistic regression model will do well by this criteria. I have provided both of these goodness-of-fit measures for the estimated logistic regression model.

3.5. Principal Components

From reading chapter 2, one notices there are many variables that could possibly affect a timber sale auction. As stated above, not all of the variables were considered as possible variables in the final model. Many of these variables were used for descriptive purposes only. However, the number of explanatory variables that were used to specify the model

approached 20. Given the large number of variables considered and that many of them are correlated, the technique of principal components was used. Principal components can achieve three goals:

1. variable reduction;
2. create new orthogonal variates that can be used in further analysis;
3. analytically determine how these variables are related.

The general form of the principal components model is:

$$Y_i = \mathbf{l}'_i \mathbf{x}$$

where Y is the new variable, the principal component, \mathbf{l}'_i is a row vector of coefficients and \mathbf{X} is the matrix of original variables. The \mathbf{l}'_i 's are chosen such that \mathbf{l}'_1 maximizes $\text{Var}(Y_1)$ and \mathbf{l}'_2 maximizes $\text{Var}(Y_2)$ subject to $\mathbf{l}'_1 \mathbf{l}'_2 = 0$ ($\mathbf{l}'_1 \mathbf{l}'_2 = 0$ indicates that two vectors \mathbf{l}'_1 and \mathbf{l}'_2 are orthogonal, at 90 degrees to each other). These new variates have a mean, $E(Y) = \mathbf{l}'_i \boldsymbol{\mu}$ and a variance, $\text{Var}(Y) = \mathbf{l}'_i \boldsymbol{\Sigma} \mathbf{l}_i = \lambda_i$ (Johnson and Wichern, 1982). λ_i is the i th eigenvalue, found by solving the characteristic equation,

$$|\boldsymbol{\Sigma} - \lambda \mathbf{I}| = 0$$

It can be shown that the \mathbf{l}'_i 's are nothing more than the eigenvectors which correspond to their respective eigenvalues (Johnson and Wichern, 1982; Morrison, 1976). The eigenvectors are found by solving the following equation:

$$\boldsymbol{\Sigma} \mathbf{x} = \lambda \mathbf{x}.$$

Geometrically, the eigenvectors define new axes in the plane (or of an ellipse). It is these new axes that define the principal components.

The calculation of the principal components is usually based upon the variance-covariance matrix, Σ . If the variables are measured in the same units or if the variances are equivalent, then the analysis should be based on the variance-covariance matrix. But this occurrence is rare, especially when considering the variables in this research. For example, average diameter is measured in inches and stump to mill costs is measured in dollars per thousand board feet; not only are the units quite different, but the variances will also be quite different. When this problem occurs, the analysis of principal components should be based on the correlation matrix, ρ . The variables in the correlation matrix are all unitless and have a variance equal to one. This eliminates the problems of units and unequal variances. The analysis is carried out in the same manner, except the matrix (ρ) is used instead of (Σ). There is one observation that should be made at this point. The outcome of the analysis based on ρ compared to Σ can be quite different, especially when considering the interpretation of the principal components (Y_i) (Johnson and Wichern, 1982). Given differences in units, my principal components analysis was based on the correlation matrix.

CHAPTER 4

EMPIRICAL RESULTS

Given the problem of estimating a statistical model (policy tool) to predict the outcome of a timber sale auction, two statistical models were chosen. The analysis began with the estimation of the linear discriminant function and the logistic regression model based upon the original explanatory variables. Once the 'goodness of fit' was determined, principal components were formed. The principal components were used as independent variables and the 'goodness of fit' for the two modeling approaches was determined. The final estimated models will be presented in the following order: linear discriminant model, logistic regression model and both of these models based on the principal components.

4.1. LINEAR DISCRIMINANT FUNCTION RESULTS

When viewing the linear discriminant function coefficients in Table 4.1 one must keep in mind how the discriminant function operates. The model produces a discriminant score that is a linear combination of the variables. When the score is less than zero the offering is predicted to be a sold offering, and when greater than zero the offering is predicted to be unsold. The interpretation of the coefficient's signs is now clear in light of how the discriminant analysis performs its classification. Variables with positive signs increase the discriminant score making it more likely that the offering will be unsold. Variables

with negative signs decrease the discriminant score increasing the likelihood that the offering will be sold.

Table 4.1:

Linear Discriminant Model		
Variable	Coefficient	Standardized Coefficient
Sale Characteristics		
1. Ln(Stump to Mill Costs)	3.307	.830
2. Haul Ratio	-.744	-.173
3. Dead	.023	.270
4. (% volume cable yarded) ²	.0001	.487
5. (Volume/Acre Harvested) ^{1/2}	-.113	-.150
6. Ln(Dense Pack)	-.183	-.148
7. (Acres Harvested) ^{1/2}	.114	.137
Administrative Characteristics		
8. Pricing Method	-.593	-.218
Economic Characteristics		
9. Ln(SPLT)	-.824	-.194
10. Uncut Volume Under Contract _{t-3}	.0006	.149
11. Exchange Rate _{t-3}	6.856	.419
12. Mill 1	-.636	-.318
Constant	-23.213	

Model significance: $x^2 = 127.80$, $DF = 12$, $P(x^2 > 127.80) = .0000$

4.1.1. SALE CHARACTERISTICS

Certain variables have a positive influence (increasing the discriminant score) in determining whether an offering will sell. The natural logarithm of stump to mill costs, percent volume dead lodgepole or dead white pine, percent volume cable yarded squared and the square root of the total acres harvested will increase the discriminant score. If these factors are present, or are increased, then it is more likely the

offering will not sell. Other variables have a negative effect (decreasing the discriminant score) in determining whether an offering will sell. The haul ratio, the square root of the volume per acre harvested, and the natural logarithm of dense pack, are factors which work to decrease the discriminant score, making it more likely the offering will sell. All signs were consistent with the proposed hypotheses.

The standardized coefficients indicate that stump to mill costs are by far the most important discriminator, with a standardized coefficient equal to .830. The percent of the volume cable yarded is the second most important discriminator, with a standardized coefficient equal to .487.

4.1.2. ADMINISTRATIVE CHARACTERISTICS

The only variable in this category, pricing method, indicates if the bid price is allowed to fluctuate with the WWPA price index it is more likely the offering will sell. This was consistent with the stated hypothesis in Chapter 2.

4.1.3. ECONOMIC CHARACTERISTICS

The following economic characteristics increase the discriminant score, causing one to predict that the offering will not sell. The uncut volume under contract lagged three months, and the Canadian/United States exchange rate lagged three months results in an increase in the discriminant score. This indicates that when these factors are increasing it is less likely the offering will sell. The remaining economic characteristics, logarithm of the selling value, lumber tally, and Mill 1, decrease the discriminant score. Thus, as the species mix is of higher

value and the Forest Service appraisal point is considered competitive, it will be more likely the offering will sell. The estimated signs were consistent with the hypothesized signs of Chapter 2.

The standardized coefficients indicate that, exchange rate_{1,3}, is the third most important discriminator. With a standardized coefficient of .419, it is slightly less important than percent of the volume cable yarded. However, it has approximately half of the discriminating power of stump to mill costs.

4.1.4. GOODNESS OF FIT RESULTS

Table 4.2 presents one of the goodness of fit measures used to quantify the linear discriminant model. The linear discriminant model correctly classified 77.9% of the timber sales. 79.3% (115 out of 145) of the unsold sales were correctly classified and 76.7% (157 out of 204) of the sold sales were correctly classified. Also, the model was extremely significant; the calculated $\chi^2 = 127.80$ with an associated probability equal to .0000. This indicated that the model using the above variables was successful at separating the two populations.

TABLE 4.2:

DISCRIMINANT MODEL CLASSIFICATION RESULTS

<u>ACTUAL GROUPS</u>	<u># OF CASES</u>	<u>PREDICTED GROUP</u>	
		<u>UNSOLD</u>	<u>SOLD</u>
Unsold Sales	145	115 (79.3%)	30
Sold Sales	204	47	157 (76.7%)
Percent of cases correctly classified: 77.9%			

Table 4.3 presents the jackknife classification results. The jackknife results are a means of verifying the stability of the classification results found in Table 4.2. From the standpoint of unsold sales, the correct classification decreases by 1 (from 115 to 114). In terms of sold sales, a 4 unit decrease is experienced (from 157 to 153). Overall correct classification decreases by 1.4% (272 to 267). These results indicate that the model is stable and that the error rate presented in Table 4.2 is a good estimator of the true error rate.

TABLE 4.3:

DISCRIMINANT MODEL JACKKNIFE CLASSIFICATION RESULTS

<u>ACTUAL GROUPS</u>	<u># OF CASES</u>	<u>PREDICTED GROUP</u>	
		<u>UNSOLD</u>	<u>SOLD</u>
Unsold Sales	145	114 (78.6%)	31
Sold Sales	204	51	153 (75.0%)
Percent of cases correctly classified: 76.5%			

4.2. LOGISTIC REGRESSION RESULTS

The interpretation of this model is straightforward, as long as we keep in mind that the parameter estimates are a change in the log of the odds that an offering will sell given a unit change in the independent variables. The logistic regression model is presented in Table 4.4.

Table 4.4:

Logistic Regression Model

Variable	Coefficient	Std Error	Coef/S.E.
Sale Characteristics			
1. Ln(Stump to Mill Costs) _{t-2}	-2.654	.357	-7.434
2. (% volume cable yarded) _{t-2}	-.00009	.00002	-4.500
3. Dead	-.018	.006	-2.779
4. Ln(Haul Ratio)	.323	.124	2.604
Administrative Characteristics			
5. Pricing Method	.462	.203	2.276
Economic Characteristics			
6. (SPLT) _{t-2}	.00001	.000004	2.237
7. MILL 1	.432	.144	3.000
8. Ln(Uncut Volume) _{t-3}	-4.670	1.816	-2.572
9. Lumber Production _{t-3}	.013	.005	2.600
10. Exchange Rate _{t-3}	-2.982	1.562	-1.909
Intercept	60.235	14.221	4.236

Goodness of fit statistics: Quadratic Score = .675.

4.2.1. SALE CHARACTERISTICS

The natural logarithm of stump to mill costs, the percent volume cable yarded squared, and the percent volume dead lodgepole or dead white pine decrease the log of the odds that an offering will sell. If stump to mill costs are high, or the sale has some form of cable yarding, or the species composition consists of dead lodgepole or dead white pine the probability of the sale selling will be reduced by the amount of their estimated coefficient and the given probability level. The last variable in this category, the natural logarithm of haul ratio, has a positive effect on the probability of the offering selling. When two competing milling centers are found close to the sale area, the

probability of the offering selling increases. All of the signs were consistent with the stated hypotheses.

4.2.2. ADMINISTRATIVE CHARACTERISTICS

If the escalation clause was used in order to establish the value of the stumpage (pricing method), the probability of the offering selling increases. This agreed with the hypothesis stated in Chapter 2.

4.2.3. ECONOMIC CHARACTERISTICS

If the appraisal point is competitive (mill 1), if the species composition is of high value, or the lumber production in the Intermountain Zone is increasing, holding all other things constant, the probability of the offering selling will increase. The reverse is true when considering the natural logarithm of uncut volume under contract lagged three months and the Canadian/United States exchange rate lagged three months. If these measures are increasing, then the probability of the sale selling will decrease holding all other things constant. The estimated signs were consistent with the hypotheses set forth in Chapter 2.

4.2.4. GOODNESS OF FIT RESULTS

The logistic regression model when used as a classification tool correctly classified approximately 77% of the sales (see Table 4.5). The model achieved only 69.0% (100 out of 145) correct classification when considering unsold sales, but correctly classified 83.3% (170 out of 204) of the sold sales. The quadratic score (average score = .675) indicated that the model fits well; recall a score close to one indicates a good fit and a score close to -1 indicates a poor fit. Both

techniques used to measure goodness of fit, number correctly classified and quadratic score, indicate a good fit.

Both estimated models proved to be adequate predictive tools. The logistic regression model and the discriminant model correctly classified 77.4% and 77.9% respectively. The dissimilarity occurred when considering the unsold sales; the discriminant model correctly classified 79.3% in comparison to 69.0% for the logistic regression model.

TABLE 4.5:

Logistic Regression Classification Results*

<u>Actual Groups</u>	<u># of cases</u>	<u>Predicted Groups</u>	
		<u>Unsold</u>	<u>Sold</u>
Unsold Sales	145	100 (69.0%)	45
Sold Sales	204	34	170 (83.3%)
Percent of cases correctly classified: 77.4%			

* Decision rule: Estimated probability $\geq .50$, then offering is sold.
Estimated probability $< .50$, then offering is not sold

4.3. PRINCIPAL COMPONENTS MODEL

Many of the variables presented in Chapter 2 are interrelated. For example, the miles of road construction will be related to total volume harvested and the contract length. Given these interrelationships, principal components were estimated and used in the modeling process. All the variables from Chapter 2 were used in the principal components analysis. The resulting principal components are linear combinations of the original variables, and thus, the interpretation of the principal

components can be difficult. I did not attempt to interpret the principal components; only the predictive results using the principal components will be presented.

4.3.1. LINEAR DISCRIMINANT FUNCTION

Table 4.6 presents the estimated coefficients and the classification results. Comparing the classification results found in Table 4.6 and Table 4.2 indicates that there is only a slight difference in the classification results. Using the principal components the percent correctly classified has increased to 78.2% (273 out of 349) from 77.9% (272 out of 349). The principal components did not capture any "new" information not already found in the original variables used in the model.

Table 4.6:

Discriminant Model Using Principal Components

Variable	Coefficient
1. Principal Component 2	.414
2. Principal Component 5	-.649
3. Principal Component 7	.279
4. Principal Component 8	-.457
5. Principal Component 13	.439
6. Principal Component 19	-.231
7. Principal Component 20	.380
8. Principal Component 21	-.175
9. Principal Component 31	.229
10. Principal Component 32	.202
Intercept	.000

Classification Results

<u>Actual Groups</u>	<u># of cases</u>	<u>Predicted Groups</u>	
		<u>Unsold</u>	<u>Sold</u>
Unsold Sales	145	118 (81.4%)	27
Sold Sales	204	49	155 (76.0%)

Percent of cases correctly classified: 78.2%

Model Significance: $x^2 = 115.226$, $DF = 10$, Significance = .0000

4.3.2. LOGISTIC REGRESSION RESULTS

Table 4.7 presents the estimated coefficients and the classification results. There was no change in the model's predictive ability. The principal components model correctly classified 77.4% of the sales (270 out of 349), which was identical to the logistic regression model using the original untransformed variables.

Table 4.7:

Logistic Regression Model using Principal Components

Variable	Coefficient	Std. Error	Coeff/S.E.
1. Principal Component 2	.258	.068	3.784
2. Principal Component 5	-.407	.071	-5.664
3. Principal Component 7	.207	.070	2.941
4. Principal Component 8	-.306	.073	-4.156
5. Principal Component 13	.267	.069	3.864
6. Principal Component 19	-.149	.065	-2.284
7. Principal Component 20	.247	.067	3.634
8. Principal Component 21	-.115	.069	-1.659
9. Principal Component 31	.153	.069	2.217
10. Principal Component 32	.147	.066	2.214
Intercept	5.252	.067	78.415

Classification Results

<u>Actual Groups</u>	<u># of cases</u>	<u>Predicted Groups</u>	
		<u>Unsold</u>	<u>Sold</u>
Unsold Sales	145	102 (70.3%)	43
Sold Sales	204	36	168 (82.8%)
Percent of cases correctly classified: 77.4%			

In general, these results indicate that the modeling process was not improved by using principal components. In fact, the results were much more difficult to interpret, since the "new" variables are a linear combination of all variables, and if basic assumptions of the model are not met it is difficult to rectify the problem. Granted, theoretically only certain variables are highly correlated with the principal component, but this is not guaranteed. If latent variables could have been observed from this modeling process, not only would this help the modeling process, but would help describe the underlying relationships to the decision maker.

In summary, the modeling process was successful in producing a model that predicts if an offering will sell given sale characteristics, administrative characteristics and economic characteristics. The principal components did not aid the modeling, both from a prediction and interpretability standpoint.

CHAPTER 5

CONCLUSIONS

Sold and unsold timber offerings became a common occurrence during the period of 1980 through 1985 in Region 1. The large number of unsold timber offerings had many forest managers puzzled as to the problems with the offerings and what could be done to better understand the offerings that were planned and ready for auction. This was not a study that would investigate Forest Service timber sale policy; the objective of the study was to formulate the problem of unsold timber offerings into a statistical model that could be used to provide planning information. The statistical model (planning tool) could provide important information to the planner concerning such things as, given the sale offering, will it sell in today's market, and what characteristics are statistically important in determining a sold offering or an unsold offering.

Two statistical models (Discriminant Analysis and Logistic Regression) were chosen in order to address the problem. Both models allowed us to model a yes/no occurrence (sold/unsold). Sale characteristics, administrative characteristics and market characteristics were used as the information in order to specify the models.

Both models, the discriminant function and the logistic regression, were successful at predicting sold and unsold timber offerings. They

correctly classified approximately 80% of the offerings. The models were estimated using various types of information, sale characteristics, administrative characteristics, and economic characteristics, all of which are available to the planner before the timber sale auction. In the beginning we assumed that this process could be modeled using information that forest planners have available to them before the auction. The models point to the fact that the planner can begin to develop a means of measuring the sale potential in an efficient fashion using available data.

If the models are straightforward and use information that is available to forest planners before an auction, these models could be used by planners during the timber sale planning process. Computerized timber sale planning tools are currently used on certain forests. These tools allow the planner to view how costs are affected given different planning scenarios. If a planner thinks that certain cutting units should be cable yarded instead of tractor yarded, he will then be able to see the effect on stump to mill costs. The models estimated in this research could be integrated into the computer software which would allow the planner to view the predicted outcome of the timber sale auction, given his proposed changes. He may be able to include more cable yarding in the sale given that he knows the effect of the proposed change on the predicted timber sale auction. This allows the planner to view how changes on the ground affect the predicted outcome of the timber sale auction.

Forest managers are creating a good that will be sold in a competitive market. The manager should try, if possible, to produce a product that will sell and return the most money possible to the United States Treasury. Given models like the ones estimated here, the forester can develop sales so as to avoid the occurrence of unsold sales. The manager should be thinking in terms of sale monitoring, and planning tools like the ones developed here will give the planner the ability to do this. However, one must remember that the manager does not control the cycles of the economy, and during downturns in the market, sales that have the characteristics of a viable sale, may not sell. This is not a product problem; it is a demand problem.

Given that these equations are stochastic models and that the timber market is dynamic, the models presented are not the final models. The equations would have to be updated (re-estimated) in order to account for changes and may even require reworking to reflect major changes in the industry.

5.1. Further Research

Several promising areas of research related to this topic exist. They deal with concepts of market areas, log markets and sale monitoring.

The first concerns market areas. The model was built using data from seven forests on the west side of Region 1. This area is very diverse when considering terrain, timber types, milling centers and potential buyers. Research could be conducted in establishing market areas within these seven forests. For example, the market area

definition could be based on log flows within and between different forests. But the forest level definition may be too large. Forests could be divided into their districts and then aggregated into market areas. For example, the Nezperce National Forest could be grouped with the two southern districts of the Clearwater National Forest. This would reflect logs flowing to mills in the Grangeville, Idaho area.

The second area of research concerns the transition of the market to a log market. Today more mills are relying on the log market to supply their log needs. Years ago the large purchaser, like Champion International, bought and logged Forest Service timber sales. Mills are now purchasing much of their needed input in the form of delivered logs. The Forest Service still relies on pricing the logs from a sale in the form of lumber output, selling price, lumber tally. But the small purchaser does not have a mill and is at the mercy of what the mill is paying for different species of logs. The models presented could be specified to reflect this transition to a log market by using a weighted average delivered log price instead of selling price, lumber tally. However, an accurate time series doesn't exist.

The final area of research concerns expanding the notion of sale monitoring. The sold and unsold classes of sales are not the only ones of interest. The classes could be expanded to include noncompetitive sales. There are several ways to measure a noncompetitive sale, but for this discussion we can consider noncompetitiveness to be a sale that receives only one bid. In this work, noncompetitive sales were within the class of sold sales. The models could be expanded to include this

additional class rather easily. If this is done, the model would be a more complete sale monitoring tool.

BIBLIOGRAPHY

- Ameniya, T. "Qualitative Response Models: A Survey" Journal of Economic Literature. vol.19, pp.1483-1536.
- Branman, L. Klein, J.D., and Weiss, L.W. "The Price Effects of Increased Competition in Auction Markets" Review of Economics and Statistics. vol. 1, pp.24-32.
- Buongiorno, J. and Young, T. "Statistical Appraisal of Timber with an Application to the Chequamegon National Forest" Northern Journal of Applied Forestry. vol. 1, pp.72-76.
- Connaughton, K.P. "Stumpage Valuation for National Forest Planning under Single and Multiple Bidder Timber Sales" Unpublished study plan, pp.1-27.
- Eisenbeis, R.A. "Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics" Journal of Finance. vol. 32, no.3, pp.875-900.
- Haynes, R.W. "Competition for National Forest Timber in the Northern, Pacific, Southwest, and Pacific Northwest Regions. USDA Forest Service Research Paper, PNW-266, 1980.
- Huang, F.M. and Buongiorno, J. "Market Value of Timber when some Offerings are Not Sold: Implications for Appraisal and Demand Analysis" Forest Science. vol. 32, no.4, pp.845-854.
- Jackson, D.H. "Why Stumpage Prices Differ Between Ownerships: A Statistical Examination of State and Forest Service Sales in Montana" Forest Ecology and Management. vol. 18, pp.219-236.
- Jackson, D.H. "An Integrated Approach to Defining Operable Timber Stocks" Proceedings of a Symposium held March 14, 1983, Missoula, Montana.
- Jackson, D.H. and McQuillan, A.G. "A Technique for Estimating Timber Value Based on Tree Size, Management Variables, and Market Conditions" Forest Science. vol.25, no.4, pp.620-626.
- Johnson, R.N. "Oral Auction versus Sealed Bids: An Empirical Investigation" Natural Resource Journal. vol.19, no.1, pp.315-335.
- Johnson-True, C. 1985. An Analysis of Forest Service Stumpage Prices in Western Montana: Prices Related to Sale Location and the Marketplace. Missoula: University of Montana. 44 p. Thesis.

- Johnson, R.A. and Wichern, D.W. Applied Multivariate Statistical Analysis. Englewood Cliffs, N.J.: Prentice-Hall. 1982.
- Koutsoyiannis, A. Theory of Econometrics. Hong Kong: Macmillan. 1977.
- Lachenbruch, P.A. Discriminant Analysis. New York, N.Y.: Hafner. 1975.
- Lyon, L.J. et al. Coordinating Elk and Timber Management (Final Report of the Montana Cooperative Elk-Logging Study 1970-1985, January 1985, pp.1-53.)
- Matthews, D.W. Cost Control in the Logging Industry. First Edition. New York: McGraw-Hill. 1942.
- McCarl, B.A. and Haynes, R.W. "Exchange Rates Influence Softwood Lumber Trade" Journal of Forestry. vol.83, no.6, pp.368-370.
- Mead, W.J. et al. The Effectiveness of Competition and Appraisals in the Auction Markets for National Forest Timber in the Pacific Northwest (Final Report, USFS, September 1981, pp.1-328.)
- Merzenich, J.P. "Developing Timber Costs and Prices for FORPLAN" Mimeo. USDA Forest Service, Region 1, Missoula, Mont.
- Milgrom, P.R. and Weber, R.J. "A Theory of Auctions and Competitive Bidding" Econometrica. vol. 50, no. 5, pp.1089-1122.
- Morrison, D.F. Multivariate Statistical Methods. New York, N.Y.: McGraw-Hill. 1976.
- Pindyck, R.S. and Rubinfeld, D.L. Econometric Models and Economic Forecasts. New York, N.Y.: McGraw. 1981.
- Press, S.J. and Wilson, S. "Choosing Between Logistic Regression and Discriminant Analysis" Journal of American Statistical Association. vol. 73, no. 364, pp.669-704.
- Winkler and Murphey "Good Probability Assessors" Journal of Applied Meteorology. vol. 7, pp.751-758.