AN ABSTRACT OF THE THESIS OF

Mark Edward Thompson for the degree of Master of Science
in Agricultural and Resource Economics presented on June 5, 1980
Title: Estimating Residential Flood Control Benefits Using Implicit Price Equations

Abstract approved:
Herbert H. Stoevener

The objective of this study was to use implicit price equations to estimate residential benefits from flood control on Sutherlin Creek, Oregon. To accomplish this it was necessary to regress land values on several price determining characteristics which included a measure of flood hazard. Two flood hazard variables, elevation above mean sea level and an indicator variable depicting a flood plain location, were considered for use in the implicit price analysis. The indicator variable was found to be superior to the elevation variable in measuring the flood hazard.

To estimate flood control benefits it was necessary to estimate the expected damages with and without flood control. Damages without the project were estimated by the coefficient on the flood plain indicator variable in a cross-sectional regression model using 1962 data. The flood control provisions were installed five years later. Remaining flood damages were also estimated with the flood plain indicator coefficient in a cross-sectional model using 1978 data.

Total residential flood control benefits were estimated by calculating the estimated reduction in damages per residential lot and multiplying this figure by the number of lots in the flood plain. This calculation in-

volved inflating the 1962 estimate of damages per lot to reflect the increase in residential land values in the Sutherlin area between 1962 and 1978 and then subtracting the 1978 estimate of damages per lot. Using the above procedure it was found that damages per lot are less with the project than they would have been without the project.

Estimating Residential Flood Control Benefits Using Implicit Price Equations

by

Mark Edward Thompson

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed June 1980

Commencement June 1981

APPROVED:

Professor of Agricultural and Resource Economics
in charge of major

Head of Agricultural and Resource Economics

Dean of Graduate School

Date thesis is presented June 5, 1980

Typed by Dodi Snippen for Mark Edward Thompson

ACKNOWLEDGEMENTS

The author is indebted to many persons for advice and assistance in the preparation of this thesis. In particular, debts of gratitude are due to:

Dr. Herbert H. Stoevener for his guidance and consultation as a major professor and for his stimulating insights on the research in general.

Drs. Jack Edwards, Roger Kraynick, and Ron Oliveira for their helpful suggestions and comments at various stages of this research project.

The personnel at the Douglas County Assessor's Office for their cooperation in the search for appraisal data.

The personnel at Sutherlin City Hall, particularly Lloyd Norris, for their cooperation in providing materials necessary for the quantification of several land price determining characteristics.

Dodi Snippen for her time spent typing and proofing the various drafts.

Finally, Cyndi for all her patience and understanding during the completion of this project.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

ESTIMATING RESIDENTIAL FLOOD CONTROL BENEFITS USING IMPLICIT PRICE EQUATIONS

CHAPTER I

INTRODUCTION

The Problem

The rivers and streams of the United States provide a diverse set of services which includes transportation, irrigation, recreation, and of course, industrial and domestic water uses. Several types of economic activities locate in areas near rivers and streams to take advantage of these services. However, the tendency for rivers and streams to exceed their banks has resulted in heavy damages to activities located within flood plains.

Our federal government has adopted a policy of altering the stream flow characteristics of rivers to reduce the possibility of flooding. Public involvement in the provision of flood control has a long history. Although some programs date back to the late 1800's the Flood Control Act of 1936 is recognized as the act which made flood control the responsibility of the federal government. The 1936 act designated the U.S. Army Corps of Engineers the responsbility of developing a large-scale flood control program. Subsequent legislation, the Watershed Protection and Flood Prevention Act of 1954, broadened the scope of federal involvement with flood control to include the smaller tributaries upstream from the large rivers.

The costs of flood control paid by the federal government have been substantial. The current administration's tightening of federal expenditures in the water resource area indicate that there may be some desire to reassess the rationale for flood control expenditures. Accurate methods for quantifying the benefits derived from flood control provisions must be available for this purpose.

Quantification of the benefits of residential flood control has largely been accomplished through the use of damage-frequency relationships. This approach requires large amounts of hydrological and economic data which often times does not exist. Another method for quantifying residential flood control benefits is the comparable area approach. By comparing land values in areas similar in all respects except the degree of flood hazard the difference in land values may be observed. This difference in land value is assumed to be equal to the discounted flow of annual flood damages. However, the characteristics which determine land values are never present at the same levels between two areas of land. These differences in characteristics other than the flood hazard will be reflected in the land value in the same manner as differences in the flood hazard. The comparable area approach will, therefore, be influenced by factors other than the flood hazard.

Another method of determining residential flood control benefits is the ~~use of the damage-frequency approach which is the~~ use of land values to estimate implicit price equations. Implicit (also called hedonic) price functions have an advantage over the comparable area approach in that the effect on land value of major price determining characteristics is isolated. When other characteristics are isolated the effect of the flood hazard on land price may be determined.

The Objective

The objective of this study is to use the implicit price equation approach to estimate the benefits from flood control in Sutherlin, Oregon. To accomplish this an economic model of land prices with emphasis on the flood hazard a parcel is exposed to is developed. The model will take the form of an equation of implicit prices paid for the characteristics associated with parcels of land. A regression analysis will then be conducted using data from residential parcels in Sutherlin, Oregon, where flooding was a problem until structural flood control measures were installed. From this analysis the form of the implicit price equation unique to Sutherlin, Oregon, and the value of parameters in the equation will be estimated. An estimate of residential flood control benefits will be presented using the information generated in the regression runs.

CHAPTER II

A THEORETICAL FRAMEWORK

Introduction

The first section of this chapter is a discussion of the empirical tools available to estimate flood control benefits. The advantages and disadvantages of employing an implicit price model to quantify these damages are briefly outlined. Next, a review of the literature relevant to the formulation of an implicit price model is presented. In the third and final section of this chapter an economic model relating the benefits associated with a reduction in the flood hazard to the price of residential parcels is developed. This model is referred to as an implicit (hedonic) price function.

Arguments For Using Land Markets to Ascertain the Reduction in Annual Residential Flood Damages

This section will present three possible methods for determining the reduction in annual residential flood damages, and, therefore, the benefits from structural flood control. These methods are the damage-frequency approach, the comparative area approach, and the implicit price equation approach. Public good characteristics associated with the provision of flood control do not allow the willingness to pay for it to be observed in the market place. A rational individual would, however, be willing to pay an annual fee for flood control equal to the expected value of annual damages avoided. The methods to be discussed are based on the premise that reduced damages are the correct measure of flood control benefits.

Damage-Frequency Approach

The annual expected ~~benefits~~ damages caused by flooding are dependent on the frequency with which the stream in question overflows its banks with and without flood protection and the damages sustained during each flooding event. This relationship between damages caused by a flooding event and the probability of actually sustaining these damages is summarized in what is called a damage-frequency curve. The area under the damage-frequency curve is the expected annual damages from flooding. Such a curve is constructed for both the with and without project states of nature so that the reduction in annual damages may be estimated.

The main limitation of the damage-frequency method is that its application is only possible in cases where an extensive historical record of flooding events and the resulting damages exists. The gauging history of the stream or river in question may not be long enough to derive reliable estimates of the probability of floods of various magnitudes.[1] The damages caused by floods of several magnitudes must also be estimated. This type of data requires that detailed surveys be made immediately after floods of varying intensity. When the above types of hydrological and economic data are available the damage-frequency approach is the preferred method for estimating flood control benefits. However, the required data are in most instances not available.

Comparable Area Approach

The comparable area method involves choosing an area of land similar

---

[1] The probability of a flood is expressed as the number of years between occurrences (e.g., a 50 year flood happens on the average once every 50 years). The magnitude of a flood is measured in terms of the discharge of water at flood stage expressed in cubic feet per second (cfs).

in all respects to the flood prone land except that the similar area is not exposed to any flood hazard. The land value differential between the two areas is the present value of the future flood damages. Total damages may be expressed in annual terms by applying the appropriate rate of discount. The problem with this approach is especially obvious to researchers who have used, or attempted to use it. Finding a similar area as defined above is exceedingly difficult and any dissimilarity between residential areas will be attributed to the flood hazard.

The Implicit-Price Equation Approach

The method adopted in this paper, implicit price equations, is a variation of the comparable area approach to estimating flood control benefits. The use of implicit price equations provides better estimates of flood damages because the influence on land price of characteristics other than the flood hazard are isolated so that non-comparability problems between areas are greatly reduced. However, this approach is considered much less direct than the damage-frequency approach because the markets for land must be relied on to accurately reflect the damages due to flooding. To the extent that individuals inaccurately perceive the nature of the flood threat the land markets will inaccurately reflect the actual damages. Land markets will, however, reflect the non-physical damages caused by flooding which would not be reflected in the damage-frequency approach (e.g., the psychological effects of living with the constant threat of flooding).

It may be useful at this point to indicate how implicit (hedonic) price functions differ from regression functions where land value is regressed on a set of price determining characteristics. The difference between the two is in name only. Methodologically the implicit price function is equivalent to a land price regression equation. The recent popularity of the implicit (hedonic) price function terminology may be because the title is more descriptive than the land price regression function terminology.

A Review of the Literature

The use of implicit price functions to quantify the benefits to residential property owners from structural flood control has been rare. However, there has accumulated a sizeable volume of empirical studies which have employed implicit price equations to estimate the benefits from changes in other land value determining characteristics, particularly air quality. The main value to this study of the previous studies is that they provide guidance in selecting important characteristics to include in the model discussed in the final section of this chapter.

One study which used regression techniques to estimate the effect of flood hazard on the value of residential properties was done by Damianos and Shabman [2]. The main objective of this study was to determine the effect on residential property values of structural and non-structural flood control measures. Sales data from flood plains and non-flood plain parcels in Radford, Virginia, were used in a time series regression model. An indicator variable to distinguish between observations taken before construction of flood control provisions and observations taken after these provisions were installed was used to quantify the effect of structural flood control.

Another land price equation was estimated by Damianos and Shabman using only pre-project data to capture in the equation the effects of the flood hazard although no flood hazard variable was included in the equation. A without project price of land was then predicted by substituting mean values of post-project observations into the regression derived from pre-project data. Subtracting actual land prices from predicted resulted in a premium for protected parcels. This premium was almost of the same magnitude as the coefficient on the indicator variable in the first model.

Damianos and Shabman conclude that residential property values had been significantly appreciated due to structural flood control. They applied similar techniques to an area of Alexandria, Virginia, where zoning had been used to restrict flood plain development. No conclusion could be made about what effects the zoning ordinance had on land prices. The importance of their efforts to this study is that the two approaches used to quantify flood damages provided insights about possible land value models.

A study of the price determining characteristics of residential land in Los Angeles County was done by Brigham [1]. Using appraisal data, Brigham regressed residential lot values per square foot on accessibility, amenity, and topographical characteristics. The cross-sectional sample was obtained by extending arbitrarily selected rays from the civic center of Los Angeles to the urban fringe. Parcels located in blocks through which the ray passed were selected for use in the sample. Two rays were used to obtain two samples.

Accessibility to economic activities was measured using two variables. One of these variables, the distance to the Los Angeles central business district (CBD) was found to be significant with the correct signs in each of the two samples. Brigham also includes an accessibility variable which reflects the significance of viable economic areas not located in or around the CBD. This variable was also found to be significant.

The appraised value of buildings on the lot was used with an environmental variable set equal to the average value of the residential structures on a given block to quantify the level of amenities associated with a parcel. The value of buildings on the lot was significant in one of the samples. The environmental variable was not significant at all in one sample and very significant in the other. Brigham also found a residence located in the hills or mountains commanded a higher price than parcels located in the valleys. The Brigham study emphasizes the important role a parcel's location with respect to viable economic centers and transportation facilities such as an Interstate-highway has in determining land value.

A cross-sectional regression model was used in a study by Stull [11] in which property values were regressed on property characteristics. Price determining characteristics of several suburban communities within the Boston Standard Metropolitan Statistical Area were grouped into four mutually exclusive and exhaustive categories. These categories are physical characteristics of the lot and improvements, accessibility characteristics, environmental characteristics, and public sector characteristics. The central hypothesis of this study was that property values were significantly affected by their proximity to non-residential land uses.

Stull was not concerned with the effect of flood hazards on property values. However, his results are of interest because they provide direction as to what variables might be included in an implicit price equation. Specifically the study found that residential lots located next to non-residential land uses were valued less than lots located in exclusively residential areas, ceteris paribus. The Stull study also provides a useful and systematic classification of characteristics.

In another study, Jack Knetsch [7] was concerned with estimating recreational benefits using land values. Sales price per acre of residential properties located in the vicinity of 11 reservoirs built by the Tennessee Valley Authority were regressed on land price determining characteristics. The resulting equation was used to predict what the value of land next to a proposed reservoir would be if the reservoir was actually constructed. The same process was used to estimate the value of the same land if the project was not constructed. The difference between the estimates could be interpreted as the discounted flow of recreational and amenity services associated with the reservoir project.

Knetsch found that the marginal effect of the reservoirs on land values was lower at greater distances from the reservoirs. This relationship was reflected by two variables: the distance from the parcel to the reservoir and the distance squared. Other land value determining characteristics which were found to be significant are: the topography of the land, the distance from urban centers, the value of improvements per acre, and the cost of development per acre. Knetsch's results demonstrate the use of land value regression models to reflect the impact of land characteristics on values.

The effects of air pollution on land values was estimated in a study by Jaksch and Stoevener [5]. The study area was Toledo, Oregon, where residential areas were exposed to varying amounts of particulate fallout from a pulp and paper plant located in the town. The objective of the study was to develop an economic model relating air pollution with residential property values in Toledo. It was hypothesized that the economic costs of air pollution would be reflected in the value of land.

Sales of residential properties were regressed on the physical properties of the lot and improvements, a measure of the economic activity of the area around the time of sale, and a measure of air pollution. The dependent variable was expressed as value of the residence in one model and value of residence per acre in another model. A significant negative relationship between the price of residences and air pollution was established in both regression models. These results support Jaksch's and Stoevener's hypothesis that the economic costs of air pollution are reflected in property values in Toledo, Oregon.

The above review of literature has dealt with only a very small portion of the literature on land value studies. The five studies outlined above have provided information which has been helpful in identifying important land price determining characteristics. Readers interested in a comprehensive review of empirical studies dealing with the determinants of residential property values and theoretical models of residential location are referred to the work of John Schefter [9].

An Economic Model

Attention is now turned to the theoretical link between land prices and the attributes associated with a parcel of land. The attribute or characteristic of particular interest to this study is the degree of flood hazard. The discussion could, however, be in terms of any other price determining characteristic of residential property.

In theory the value of land is equal to the present value of its expected future flow of net income or, in the case of home owners, net utility. Flood damages are discounted into land values when a flood plain location involves flood induced repair and/or replacement cost (e.g., new carpet must be purchased more frequently) or disutility is associated with the use of flood prone land (e.g., psychological effects of living with the constant threat of flooding).

It is expected that the value of property will decline as the flood hazard increases. That is, flood hazard is perceived by consumers as a bad. The following discussion may proceed easier if the analysis is presented in terms of a characteristic for which more is preferred to less. For example, consider a variable (E) which is a measure of the absence of a flood threat, say, flood control. As E increases the likelihood of flooding decreases and land prices are expected to increase.

Consider an individual who derives utility from the amount of E and all other goods available to him in a given time period. His utility function could be written as:

$$U = U(E, X) \tag{1}$$

where

U = the utility function of an individual

E = the degree of flood control

X = all other goods or money

Indifference curves corresponding to various levels of utility could be drawn if the functional form of equation (1) were known. These indifference curves could look like the ones drawn in Figure 1a for two levels of utility, U^0 and U'. All of the various characteristics of indifference curves

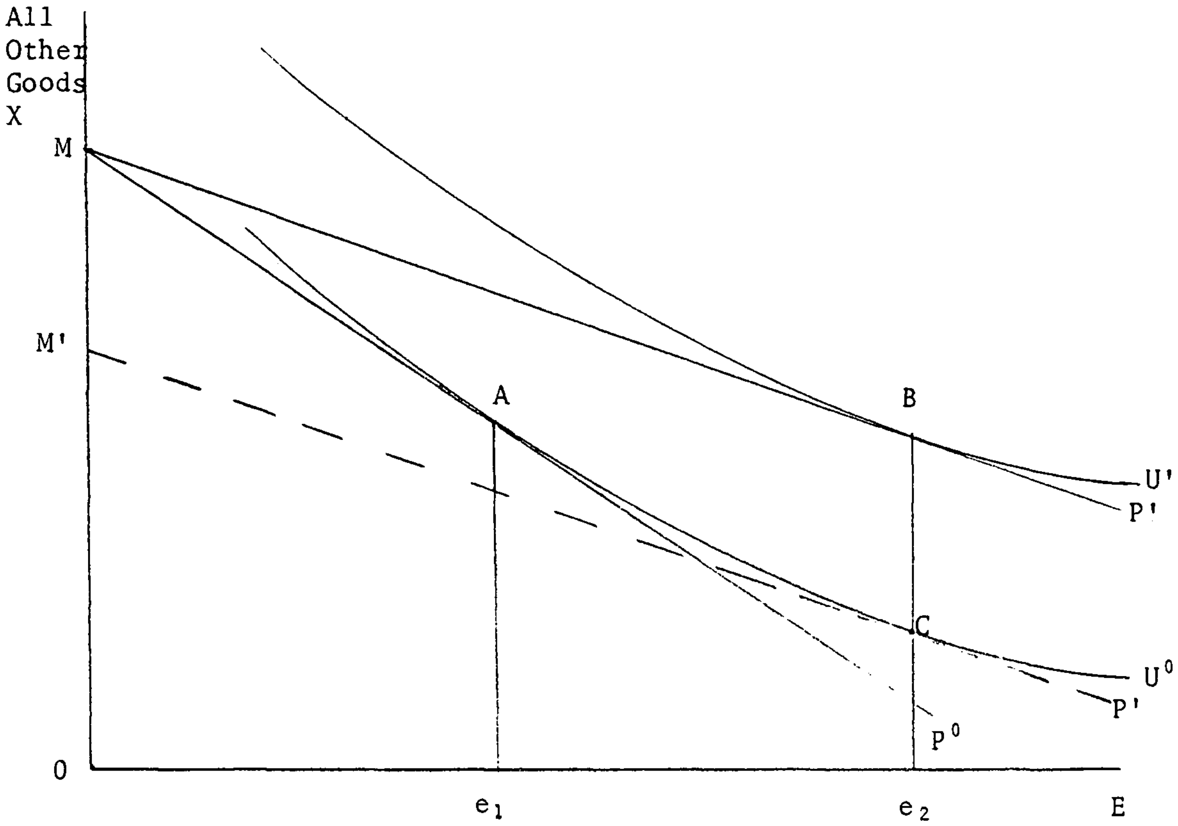


Figure 1a. Indifference Curve Analysis of Flood Control.

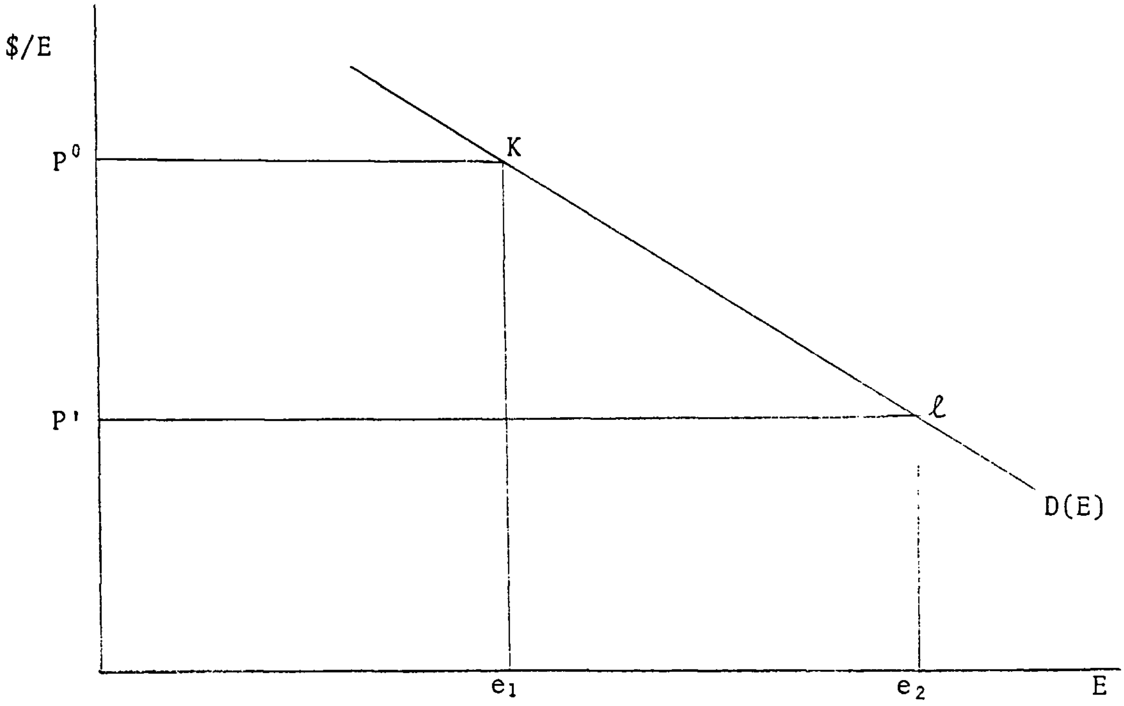


Figure 1b. Income Compensated Demand Curve for Flood Control.

are assumed to hold. The individual is equally satisfied with any combination of E and X lying on one indifference curve. No two indifference curves are allowed to cross one another. The level of utility increases as indifference curves located farther from the origin in a northwesterly direction are reached.

The slope of an indifference curve reveals how the consumer values X and E in relation to one another given levels of X and E. That is the marginal utility of X is compared to the marginal utility of E via the indifference curve. Stated somewhat differently the slope of the indifference curve reveals how much X the consumer is willing to give up for an additional unit of E keeping his utility constant. From equation (1) the willingness to trade X for E is given by:

$$\frac{\partial U/\partial X}{\partial U/\partial E} = \text{Marginal rate of substitution of E for X and the slope of the indifference curve.} \quad (2)$$

It is assumed that the consumer will attempt to maximize utility. However, he is constrained by his current level of income and the price he must pay to receive X and E. Let current income be equal to OM in Figure 1a and set the price of money equal to one. A budget constraint can then be drawn showing all possible combinations of X and E the consumer can purchase given income and prices. The slope of the budget line will be equal to the price of E. Budget lines are drawn in Figure 1a for two prices of E, P^0 and P'.

Consumer equilibrium, utility maximization, occurs when the consumer allocates his total budget between X and E in such a manner that the rate at which the market allows substitution between E for X, as given by the slope of the budget line is just equal to the rate at which the consumer is willing to substitute E for X, as indicated by the slope of the indifference curves. Stated in mathematical terms, utility is maximum when:

$$\frac{\partial U/\partial E}{\partial U/\partial X} = \frac{P_E}{P_X} = P_E \tag{3}$$

and

$$M = X + P_E E \tag{4}$$

where

M = total income of consumer

P_E = the price of E

X and E are as previously defined.

Assuming the price of E is equal to P^0 equilibrium is at A in Figure 1a with a consumer purchasing e_1 units of E. If the price of E is allowed to fall to P' the consumer will move to point B in Figure 1a. Utility has increased to U' as the amount of E the consumer will purchase increases to e_2. Assuming that the marginal utility of the numeraire commodity is constant the consumer would be willing to pay MM' to obtain the increase in utility associated with the decrease in the price of E from P^0 to P'. If the consumer is taxed MM' to extract this willingness to pay a new equilibrium will emerge shown by point C in Figure 1a. Notice that the consumer is indifferent between the equilibrium obtained with an income of OM and a price of E equal to P^0 (point A) and the equilibrium obtained with an income of OM' and a price of E equal to P' (point C).

By taxing (subsidizing) the consumer in amounts which leave him indifferent to price decreases (increases) the income compensated demand curve for E, so shown by D (E) in Figure 1b can be derived. The significance of the compensated demand curve for E lies in the fact that the benefits (willingness to pay) for increases in E (reductions in the flood hazard) may be directly obtained. For example, if structural flood control has the effect of increasing E from e_1 to e_2 the owner of that land would be willing to pay an amount of money equal to $e_1K\ell e_2$ which is also equal to MM' because of the assumption of a constant marginal utility of money.

When dealing with market goods it is usually sufficient to stop with the theory developed to this point. As Figure 1b implies, the marginal valuation of goods is equal to the price consumers pay in the market place. However, individuals are not given an opportunity to purchase the characteristics of land, such as E, independently. Hence, no prices paid explicitly for these characteristics can be observed. It was stated earlier that the price of land is likely to increase as the flood risk associated with it decreases. From this hypothesis it may be reasoned that there is an implicit price paid for E which is reflected in the total price of the property. It is, therefore, necessary to develop further the above theory to show how implicit prices for the characteristics of land are reflected in the explicit prices paid for land.

When an individual purchases a parcel of residential land he acquires a bundle of characteristics the whole of which determines his willingness to pay, or bid price, for the parcel. So, a parcel is defined by a vector of price determining characteristics, or:

$$R = (C_1, C_2, \ldots, C_n) \tag{5}$$

where

R = a unique residential parcel

C_i = the value of the i^{th} characteristic associated with R.

$i = 1, 2, \ldots, n.$

n = number of characteristics.

It is not possible to untie R and alter the measure of any one C_i. The consumer, referred to as purchaser or buyer in the remainder of this section, must take C_i as given for any one R. It is possible, however, for the purchaser to shop around and compare the C_i being offered with each R.

Before proceeding a few more assumptions must be injected. First, it must be assumed that the purchaser accurately perceives the levels of each characteristic associated with the residential properties from which he is to choose.[2/] Also, it must be assumed that there are enough residential parcels available that the buyer is faced with the continuous levels of all n characteristics. This assumption allows for discussion of the marginal conditions for utility maximization. As a result of shopping and comparing the C_i associated with various residential packages the buyer pays a price to secure the expected utility from each of the n characteristics defining the parcel of his choice. That is, the buyer reveals an implicit price function for the characteristics defined as:

$$P(R) = P(C_1, C_2, \ldots, C_n), \quad i = 1, 2, \ldots, n \tag{6}$$

[2/] This assumption should not be taken lightly, especially in the context of this study. If purchasers underestimate the flood risk associated with a parcel the theory to be discussed has no relevance to empirical studies. That is, the willingness to pay for flood free parcels will not be accurately reflected in the prices paid for residential parcels. For a discussion of the perception of exposure to risk of flooding see Kates 1962 [6]. Kates' finding indicates that the accuracy of the flood hazard perception decreases as the frequency of flooding events decreases.

where

$P(R)$ = the market price of R.

$P(C_1, \ldots, C_n)$ is the buyer's implicit (hedonic) price equation which relates prices to characteristics. It will be demonstrated that under certain conditions to be named below the first order partial derivative of P with respect to a given characteristic represents the purchasers' minimum marginal valuation of the characteristic.

Since the purchaser derives utility from more than one characteristic it is necessary to expand the utility function shown by equation (1) to:

$$U = U(C_1, C_2, \ldots, C_n, X) \qquad (7)$$

where

U, X, and C_i are as defined before. $i = 1, 2, \ldots, n$.

Set the price of X equal to one and define the amount of disposable income available to the buyer in each time period as M. The buyer will then maximize the utility function (7) by choosing a unique residential parcel R*, defined by the vector of characteristics C_i^*, subject to the budget constraint:

$$M = X + P(R) \qquad (8)$$

The first order conditions for utility maximization are:

$$\frac{\partial U/\partial C_i}{\partial U/\partial X} = \frac{\partial P}{C_i} \quad , \quad i = 1, \ldots, n \qquad (9)$$

and

$$\frac{\partial U/\partial C_i}{\partial U/\partial C_j} = \frac{\partial P/\partial C_i}{\partial P/\partial C_j}, \quad i, j = 1, 2, \ldots, n; \ i \neq j \qquad (10)$$

and

$$M = X + P(R) \qquad (11)$$

The necessary conditions for a utility maximization as expressed by equations (9) and (10) for n+1 dimensions are equivalent to the conditions expressed in equation (1) for two dimensions. Equation (9) states that the rate at which the purchaser is willing to substitute C_i for the numeraire commodity must equal the market implicit price paid for this substitution. Similarly, equation (10) states that the willingness to substitute C_i for C_j must equal the ratio of their market implicit prices. Equation (10) states that total expenditures must equal M.

A minimum value for the benefits from a decrease in the expected annual damages from flooding can be estimated using the theory of implicit prices developed above. Assume that through multivariate analysis of residential property values the price function for residential property without flood protection has been estimated and, letting E replace C_1, is of the following linear form:

$$P(R) = b_0 + b_1E + b_2C_2 + b_3C_3 + \ldots + b_nC_n \qquad (12)$$

The coefficient b_1 is the first order partial derivative of equation (12) with respect to E, and by definition is a minimum estimate of the implicit price which individuals are willing to pay to receive an additional unit of E.[3/] Because E has been defined such that more is preferred to less, b_1 is positive and represents a minimum estimate of the amount of money a consumer would be willing to pay to receive an additional unit of E and remain at his original level of utility. Total discounted damages on any given lot are equal to b_1E. The total damages associated with the flood plain are equal to:

$$b_1 \sum_{j=1}^{L} E_j,\ j = 1, 2, \ldots, L \tag{13}$$

where

L = number of residential lots on the flood plain.

In the unlikely case where a flood control project is expected to completely eliminate flood damages, equation (13) is also an expression for the benefits of the project. When damages remain with flood protection these must be subtracted from the damages without the project to estimate total benefits from flood control.

If an estimate of the annual benefits is desired it is necessary to know the duration of the flow of annual benefits stemming from the project and the rate at which these benefits are discounted. Annual benefits are equal to:

$$AB = {}^{TB}/ \sum_{i=1}^{T} \left(\frac{1}{(1+d)^T}\right) \tag{14}$$

where

[3/] The coefficient b_1 is the implicit price of E when it can be assumed that all consumers have equal incomes and utility functions. Otherwise b_1 is a minimum estimate of the marginal willingness to pay for E. For a discussion of these assumptions see Freeman [4].

AB = annual flood control benefits

TB = total flood control benefits

T = life of the flow of annual benefits in years

d = rate of discount.

Land Price Determining Characteristics

Land values are, of course, affected by characteristics other than the flood hazard. The remainder of this section will describe four broad categories of land price determining characteristics. The categories of characteristics are the same ones used by Stull [11]. They are physical, accessibility, public, and environmental characteristics.

The physical characteristics which help determine the value of a residence include improvements located on the lot and the lot itself. Physical characteristics of the improvements include the number of bedrooms and baths, the type of heating system, total floor space, etc. Physical feastures of the parcel include topography of the site, type of soil, and elevation.

The parcel's location with respect to centers of activity and transportation facilities has been labeled accessibility characteristic. The distance from major shopping centers and areas of employment are included in this category. Also, a parcel's location from important transportation facilities such as major freeways would be in this group. Accessibility characteristics have proven to be significant factors in the pricing of residential property.

Public characteristics are simply the quality and quantity of public services a community offers in relation to the tax burden the community imposes. Empirical studies of these characteristics commonly use the amount of money spent on public schools as an indication of the quality and quantity of public services offered by certain communities. Other public services which may be important land value determinants include fire protection, police protection, and community provided recreational opportunities.

Environmental characteristics refer to the social and physical aspects associated with the area where the parcel is located. The dominant ethnic group of the neighborhood where the lot is located is an example. Within this category are also included the characteristics of main interest to this study; the flood hazard asssociated with a parcel. Another environmental characteristic one may want to consider is the quality of air in the immediate proximity of the residential parcel. Damages caused by poor air quality are discounted in land value in the same manner as are flood damages.

CHAPTER III

THE EMPIRICAL PROCEDURES

Introduction

This Chapter will contain a brief description of the study area which will include a short discussion of the nature of pre-project flooding and how the boundaries of the Sutherlin Creek flood plain were determined. Also, a discussion of the characteristics thought to be important land value determinants and the economic model relevant to Sutherlin, Oregon, is included in this Chapter. This Chapter will conclude with a discussion of the data collection procedure.

The Study Area

The study area is the City of Sutherlin which has an estimated 1979 population of 4,650. The town is located in Douglas County, Oregon, approximately 75 miles from the Pacific Ocean, and 140 miles north of the California border (Figure 2). Interstate-5 which is the major highway along the west coast passes through the western edge of Sutherlin.

Major industries in the town are related to the timber industry which dominates the Douglas County economy. Two wood products plants, Georgia Pacific and Mount Mazama, operate within the city limits. Also, a machinery fabrication plant which makes parts for timber harvesting equipment is located in Sutherlin.

Sutherlin is situated 12 miles north of Roseburg which is the largest city in Douglas County. Greyhound bus lines have a terminal in Sutherlin. However, the town has no railroad freight services or commercial airlines. All of these services are available in Roseburg.

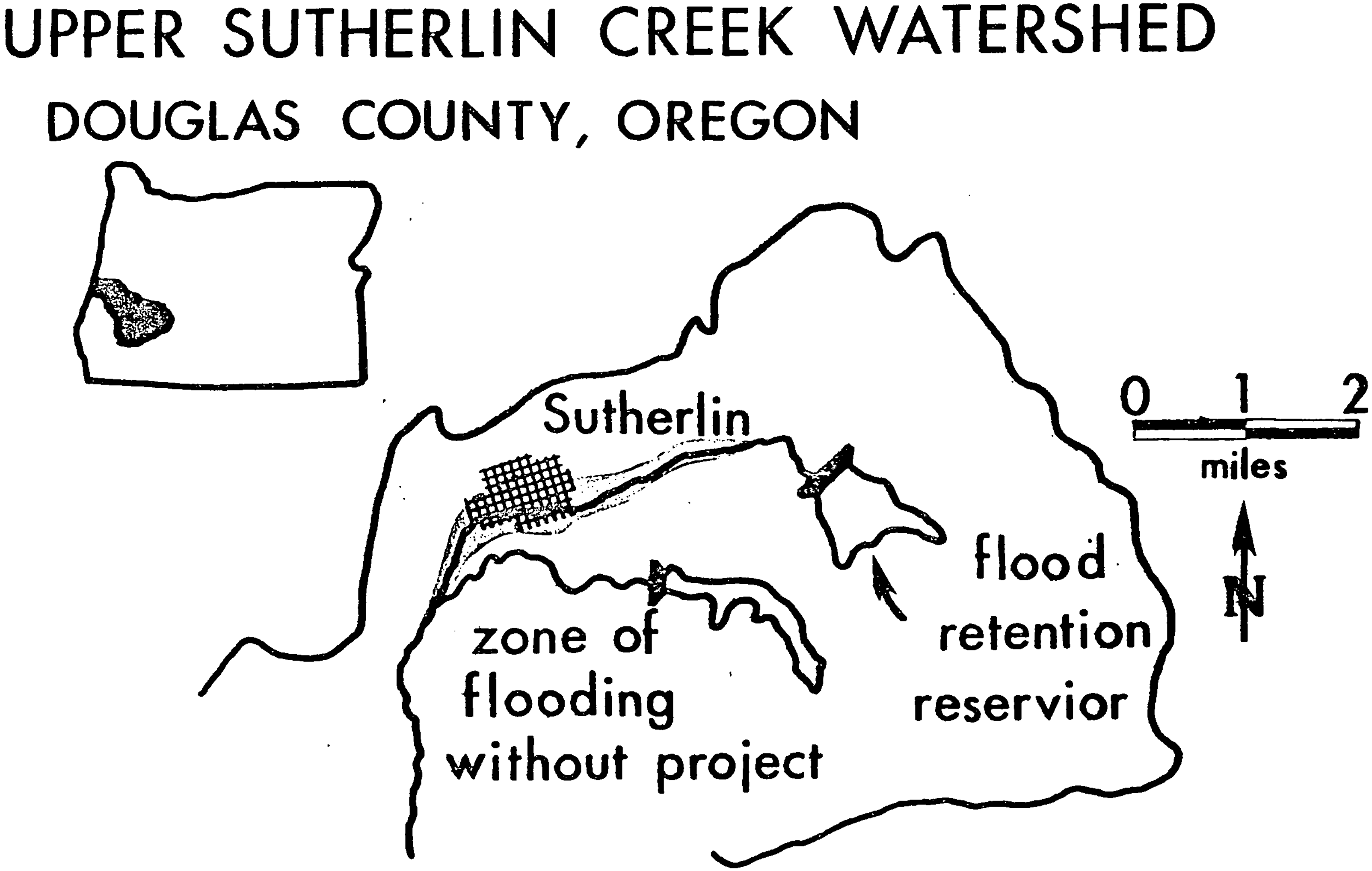


Figure 2. Location of the City of Sutherlin and the Flood Control Installations.

The Pre-Project Flood Problem and the Flood Plain

Flooding used to be a common occurrence in the Sutherlin area. Large areas of agricultural land would be inundated almost every winter as Sutherlin Creek overflowed its banks. About once in every two years water would enter the streets of residential areas and the downtown business district of Sutherlin [SCS, 1963, p. 101]. This flooding has been almost completely eliminated since installation of the flood control provisions of the Sutherlin Watershed project. There is an area of remaining flooding where Cooper Creek meets Sutherlin Creek, However, the area is very small compared to what was typically inundated before the project. Sutherlin Creek has not overflowed causing inundation of business and residential areas since the project was constructed. The works of improvement which provided for most of the flood control, Plat I Dam and the channel improvement, were completed in 1968 and 1970, respectively (Figure 2).

The flood plain is defined as that area which would be inundated without flood protection given a flood with a two percent chance of occurrence. The Sutherlin project is designed to protect against a flood of this magnitude, also called a 50 year flood. Boundaries of the 50 year flood plain were approximated using the confines of the flood which occurred February 10, 1961. The probability of a flood of that magnitude was estimated by the USGS to be six percent. This flood was used to estimate the confines of the flood plain because it was the largest flood recorded by the USGS in their 15 years of gaging efforts on Sutherlin Creek. Also, a map showing the area of inundation was available for the February 1961 flood. The difference in area inundated by Sutherlin Creek during a six percent flood and during a two percent flood is probably small due to the steep gradient around the fringe of the flood plain.

The use of Sutherlin as a study area had two principal advantages. First, the hydrological characteristics of the Sutherlin Creek watershed are such that at least moderate flooding was experienced about once every two years. Frequent reoccurrence is advantageous from the point of view of a land value study because the flood hazard and resulting damages are less likely to be inaccurately perceived. A second advantage is that the Sutherlin area, being a rather small community, is a fairly homogeneous study area.

The chief disadvantage of the study area is also related to its size. Statistical information which may have been available for larger areas is not available for the City of Sutherlin. Such statistics as property valuation by class, employment figures by industry, income parameters, and retail sales for the city would be among this type of information.

The Economic Model of Sutherlin, Oregon

In the last section of Chapter II a discussion of the general price determining characteristics of land was given. This section will discuss those characteristics thought to be relevant to the Sutherlin area. It was suggested in Chapter II that characteristics be grouped into four classes: physical characteristics of the parcel, accessibility characteristics, environmental characteristics of the area surrounding the parcel, and public sector characteristics. The discussion which follows is in terms of this classification.

Physical Characteristics

The physical characteristics of the parcel can be further subdivided into those associated with the lot and those associated with the housing unit on the lot. The size of lot has been found to be a significant determinant in previous land price studies and is concluded as a characteristic variable in this study. Land values are expected to vary directly with lot size. Physical characteristics of the improvements include total floor space, number of rooms, number of bathrooms, age of structure, condition of structure, type of plumbing, type of heating, etc. The effects of all of these characteristics taken together is considered by including only one variable, the value of the improvements. The value of a residential parcel should theoretically only be influenced by the characteristics associated with the lot and not the characteristics associated with the improvements. When a residence is sold, however, one price is paid for the bundle as a whole. To the extent that appraisal records, which are based on actual sales, do not accurately separate the value of the lot from the value of improvements the latter may be an important variable in a land value study. Land values are expected to vary directly with the value of improvements.

Accessibility Characteristics

It has been found that a parcel's location with respect to centers of employment, shopping centers, transportation facilities, and recreational facilities is a significant determinant of land price [Brigham, 1965; Knetsch, 1964]. Because of Sutherlin's proximity to the relatively large city of Roseburg it is expected that land prices in Sutherlin will be influenced by proximity to both Sutherlin and Roseburg. The distance to Sutherlin's central business district (CBD) is included as an explanatory variable on the assumption that many of the local citizens rely on Sutherlin for a significant part of their shopping needs. The effect on price of distance to Roseburg is approximated by a variable which measures the distance to a common point on highway 99 just south of Sutherlin. This variable will be called the southern access characteristic. As the distance a parcel is located from a highway has been shown to be significant in the literature on land price determinants a variable describing the distance from a lot to I-5 is included in the model as well. The coefficients on all accessibility variables are expected to be negative.

The marginal value of close proximity to the central business districts of Sutherlin, the southern access, and to Interstate-5 is expected to be higher at closer proximities and decline as one moves away from the three origins in question. To reflect this nonlinear relationship between distance and land value the squares on each accessibility variable discussed above were also included in the economic model. The coefficients on the squared accessibility variables are expected to be positive.

Environmental Characteristics

Environmental characteristics refer to the social and physical aspects of the neighborhood in which the parcel is located. Income levels, presence of industrial facilities or parks, air pollution, and flooding are among the important characteristics within this class. Three environmental characteristic variables are selected for use in the regression analysis. The average value of residential structures on the block in which a parcel is located is used to reflect income differentials between neighborhoods. Parcel values are expected to increase as the average value of improvements on the block increases. Another environmental variable is used to reflect the effect on the price of a parcel located next to a nonresidential land use. This effect is quantified by using a dummy variable equal to one when a parcel was located next to one or more of the following nonresidential land uses: commercial, industrial, schools, and churches. The expected sign of the coefficient on this variable is negative because of the congestion associated with a location next to any of the above land uses.

Within the environmental group of variables, the flooding characteristics of a parcel of land is the variable of special interest and will be used to measure the benefits of flood control to residential landowners. The effects of flooding on parcels of land will be quantified through the use of two variables in separate models. The elevation at which a flood plain lot is located will be used in some models while a dummy variable depicting location inside rather than outside of the flood plain will be used in other models. Land prices should vary directly with elevation. The coefficient on the flood plain dummy variable is expected to be less than zero indicating lower values associated with flood plain properties.

Public Sector Characteristics

Public sector characteristics could not be incorporated into this study. Given the use of cross-sectional data, there was insufficient variation in the level of public services provided to the properties in the study to describe a relationship between public services and property values.

Model Specification

This subsection will present the statistical form of the two models to be explored and evaluated in terms of their effectiveness in reflecting the hypothesized relation between land values and the flood hazard. The two models to be discussed differ in regard to the form of the independent variable used to reflect the flood hazard. Both of these models are cross-sectional. A time-series model was considered but lack of variation over time in most of the independent variables precluded use of such a model.

Before proceeding it will be helpful to define the variables as follows:

$Y_{(i)}$ = market value of lot i

$X_{1(i)}$ = Lot size of parcel (i) in square feet

$X_{2(i)}$ = Appraised value in dollars of improvements of lot (i)

$X_{3(i)}$ = Distance to Sutherlin's CBD from lot (i) in feet

$X_{4(i)}$ = Squared distance to Sutherlin's CBD from lot (i) in feet

$X_{5(i)}$ = Distance to southern access on highway 99 from lot (i) in feet

$X_{6(i)}$ = Squared distance to southern access from lot (i) in feet

$X_{7(i)}$ = Distance to I-5 from lot (i) in feet

$X_{8(i)}$ = Squared distance to I-5 from lot (i) in feet

$X_{9(i)}$ = Average appraised value in dollars of residential improvements on the block where lot (i) is located

$X_{10(i)}$ = Dummy variable equal to 1 if lot is located next to a non-residential land use

$X_{11(i)}$ = Elevation of lot (i) in feet above mean sea level (MSL)

$X_{12(i)}$ = Dummy variable equal to 1 if lot (i) is located in the flood plain and equal to zero otherwise.

Model 1

The characteristics thought to be important in determining residential land prices in the Sutherlin area, including a flood hazard variable, are included in this model. The degree of exposure to the flood threat is reflected by the elevation variable. The damages caused by flooding are expected to be less at higher elevations relative to lower elevations. There will be no damages beyond the flood plain fringe. Therefore, it is necessary to include only flood plain parcels in the estimation of the regression equation. Stated in mathematical form this model is:

$$Y_{(i)} = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \beta_7X_7 + \beta_8X_8 + \beta_9X_9 + \beta_{10}X_{10} + \beta_{11}X_{11} + \varepsilon_i \qquad (15)$$

The estimated coefficient on the elevation variable, b_{11}, when the model is run with 1962 data may be interpreted as the capitalized value of the reduction in damages associated with a one foot increase in elevation. Hence, it follows that total residential flood damages without the project are equal to:

$$\sum_{i=1}^{P} (X_{11}^{*} - X_{11_i})\, b_{11} \qquad (16)$$

where:

X_{11}^{*} = elevation of the flood plain fringe

P = the number of residential parcels located within the flood plain

X_{11_i} and b_{11} as previously defined.

The remaining damages are calculated in the same manner except for using 1978 data to estimate equation (15). If the effect of X_{11} is found to be nonsignificant in the 1978 run the flood hazard is no longer an important land value determining characteristic. Hence, remaining damages may be assumed to be close to zero.

Model 2

This model contains an alternative form of the variable used to reflect the flood threat. It focuses on the effects of the flood threat on land values by estimating the differences in value between parcels exposed to the threat and those not exposed to it. This is accomplished by including a dummy variable equal to one if the parcel is located in the flood plain and equal to zero in the case of a non-flood plain location. Observations from on and off the flood plain must be included in the analysis. Stated in mathematic form Model 2 becomes:

$$Y_{(i)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{12} X_{12} + \varepsilon_i \quad (17)$$

The only difference between this model and Model 1 is that X_{12} has replaced X_{11}. The absolute value of the estimated coefficient on the flood plain dummy variable, b_{12}, may be interpreted as the present value of the discounted stream of flood damages without the project (when Model 2 is run using 1962 data) and with the project (when Model 2 is run using 1978 data). Total residential inundation reduction benefits are equal to:

$$P \left[b_{12(1962)} - b_{12(1978)} \right] \quad (18)$$

where

P = the number of residential parcels located within the flood plain

$b_{12(1962)}$ = the total flood damage per lot without the project

$b_{12(1978)}$ = the total flood damages per lot with the project.

The non-comparability between the beta coefficients of the two years must be eliminated by either inflating $b_{12(1962)}$ (so that benefit estimates are in 1978 dollars) or deflating $b_{12(1978)}$ (so that benefit estimates are in 1962 dollars) by an index reflecting the changes in land values between 1962 and 1978.

Data Collection

In this section a brief discussion of the sampling procedure is presented. The source and method of quantification of the dependent variables as well as the independent variables is also discussed.

A total of 117 residential parcels were selected for observation. Of these, 65 were situated in the flood plain of Sutherlin Creek. Their selection proceeded as follows.

First, all blocks were identified which were situated entirely inside the flood plain. All residential lots located on blocks meeting this criterion were included in the flood plain sample. Flood plain lots were chosen in this manner because appraisal data are not likely to reflect the effects of the actual flood plain boundary. That is, all appraised values on a block with portions of lots inside and outside of the flood plain may be based on the sale of only a non-flood plain lot, for example.

Non-flood plain observations were taken from the two established residential areas of pre-project Sutherlin which were never flooded. These areas are located north and west of the central business district. Two blocks in the area north of the central business district were

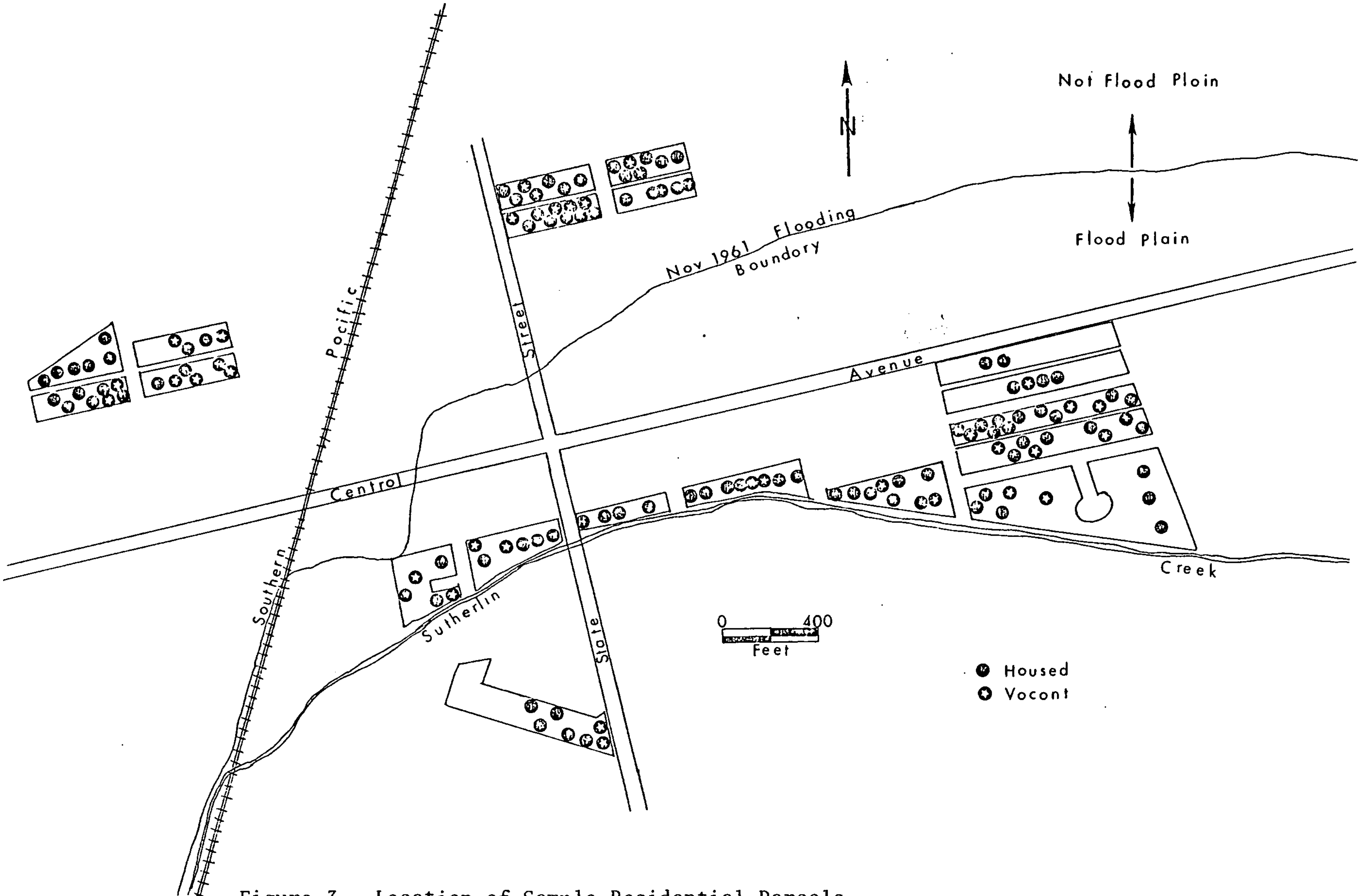


Figure 3. Location of Sample Residential Parcels.

selected for sampling along with two blocks in the area west of the central business district. The locations of the flood plain and non-flood plain sample blocks are shown in Figure 3.

It was originally hoped that actual sales data could be used to measure the dependent variable. However, the sales price of a parcel was rarely included with the information listed on the recorded deeds. An extensive search of property deeds in the Douglas County courthouse uncovered 58 sales of residential property within the Sutherlin Creek flood plain between the years 1950 and 1968, inclusive. Of these only eight deeds listed the actual price paid for the property. Therefore, it was necessary to use appraisal data so that a sufficient number of observations could be included in the statistical analysis.

The use of appraisal records is less desirable than actual sales data because the former are only an approximation of actual market values. However, the objective of the appraisal process is to estimate such values and it is believed to provide adequate proxies for parcel values.

The appraised value of parcels was recorded on individual appraisal packets made available by the Douglas County Appraiser's Office. Before 1972 properties were appraised only once every several years in the Sutherlin area. Appraisals were done in 1958, 1962, and 1968. Hence, these are the only pre-project years for which measures of the dependent variable are available. Some lots in the sample were appraised in each of the three years while others were appraised in only one or two of them. Hence, the sample size varied from year to year with 103, 59, and 107 parcels comprising the sample in years 1958, 1962, and 1968, respectively. Data on two of the independent variables were also collected using the appraisal packets.

It is necessary to select one pre-project year and one post-project year for use in the analysis. Each model will be used to quantify the present value of expected flood damages with and without the project. Therefore, the pre-project year should be the year which is closest to the year in which the project was constructed, yet not so late that construction or planning activities have begun to change the way in which individuals perceive the future hazards of locating in the flood plain.

Planning activities for the Sutherlin Creek Watershed Project started as early as 1959. However, the project was not authorized for construction until 1964 with major construction activities beginning in 1966. Based on the timing of planning and construction activities 1962 was chosen as the year which best reflects "without project" conditions. The "with project", or remaining damages, will be estimated using 1978 data. This is considered the best year because land values have had the maximum time to adjust to the lower flood hazard. Measures of size of the parcel and appraised value of improvements on the lot were included among the information in the packet.

The average value of residential improvements on a given sample block was obtained from aggregating appraisal packet data. Aerial photographs were used in conjunction with a large map of the City of Sutherlin to calculate the distance in feet to Interstate-5, the Sutherlin central business district, and the southern access. The elevation of each parcel was determined with topographic overlays and plat maps. The plat maps were also used with land use information coded on the Douglas County tax role to determine if a sample parcel was located next to a non-residential land use. A parcel was coded as a flood plain location if it was found to be within the boundaries of the February 1961 flood.[1]

---

[1] This flood appears to have been mapped using overlays on aerial photographs. The individual or organization that prepared the map is not known.

CHAPTER IV

STATISTICAL ANALYSIS

Introduction

In Chapter III two regression models were proposed for evaluating the effects on land values of exposure to the flood hazards of Sutherlin Creek. Model 1 contains an elevation variable to capture this effect while Model 2 contains an indicator variable depicting flood plain location. This chapter will present statistical results from application of both models. Based on these results one model will be chosen as the model best suited for use in this analysis. Various forms of the best model will be explored with the objective of picking the most appropriate form for 1962 and 1978. In the final section of this chapter the residential inundation reduction benefits from flood control on Sutherlin Creek will be estimated using results of the 1962 and 1978 implicit price equations.

Regression Results

Before reviewing the results of Model 1 and 2, it is useful to describe the various ways in which individual variables were expressed or the data were aggregated in the search for the statistically superior regression equation. The dependent variable was tested in two forms. In addition to the land value per lot, the land value per square foot was used in each of the two models. Also, the sample was divided into two subsamples. These consisted of vacant lots only and lots with houses (improvements) only. Experimentation with two forms of the dependent variables, two data subsamples as well as the data set combined, made it necessary to estimate four regression equations for each model. There was an insufficient number of observations for the vacant lots only subsample. The results of this experimentation will be discussed by model. The results of Model 1 will be compared with Model 2 using 1962 data. This year, for reasons discussed in Chapter III, should best reflect flood damages without the project.

Model 1 contains an elevation variable to estimate the damages caused by flooding. Table 1 lists the results of the four regression equations estimated for Model 1.[1/] The statistical significance of X_{11} is very low in all of the equations. Discussion of which equation best suits the needs of this study is not necessary since all are quite poor.

These results are perhaps not surprising when one considers that the physical features of the flood plain allow for little variation in the elevation of sampled flood parcels. The flood plain has very little gradient except around its fringe. The sampling procedure precludes the inclusion of parcels located close to the flood plain fringe for reasons discussed earlier. Furthermore, it is doubtful that appraisal data are sensitive enough to pick up land value differences resulting from small changes in elevation. The use of this variable may be enhanced by including a greater range of elevations and adjusting elevations so that they are measured as a first difference with respect to the streambed elevation along a line normal to the stream.

[1/] The squared distance variables were left out of the regression equations presented in Table 1 and those equations presented in Table 2 for Model 2. These variables, as one would expect, are highly correlated with the first order distance variables. For this reason they are only considered for inclusion in a regression equation once a selection between Model 1 and Model 2 has been made and alternative functional forms of the model are to be explored.

Table 1. Results of Regression Model 1 With Two Forms of the Dependent Variables and Two Subsamples.

Equation	Constant	X_1	X_2	X_3	X_5	X_7	X_9	X_{10}	X_{11}	R^2	N	F
Dependent Variable = Land Value Per Lot												
(1a) All lots in sample	943.2 (0.07)	0.04* (3.00)	-0.01 (1.03)	-0.09 (0.97)	0.08 (0.97)	-17.61 (1.54)	0.03 (2.76)	-52.9 (0.61)	.047 (0.02)	.548	30	3.18
(1b) Improved lots only	-1454 (0.11)	0.05* (3.24)	-0.02 (1.16)	-0.07 (0.76)	0.02 (0.21)	-10.4 (0.89)	0.03* (2.51)	-26.21 (0.27)	4.22 (0.16)	.563	26	2.74
Dependent Variable = Land Value Per Square Foot												
(1c) All lots in sample	1.17 (0.36)		-.000007* (2.10)	-.00003 (1.11)	.00002 (0.88)	-.003 (1.02)	.000006* (2.83)	-.007 (0.31)	-.002 (0.29)	.449	30	2.56
(1d) Improved lots only	-0.25 (0.11)		-.00004 (1.10)	-.00001 (0.81)	-.000004 (0.27)	-.0006 (0.28)	.000004* (2.33)	-.006 (0.33)	.0007 (0.15)	.318	26	1.20

* indicates statistical significance at the $(1 - \alpha) = .90$ level or greater.

Variables which are significant and correctly signed in at least one of the equations in Table 1 include lot size and the average value of residential improvements on the block. In addition to the measurement problems discussed above the elevation variable is also highly correlated with the accessibility variables in the equation. One possibility of improving the significance of Model 1 is to drop the accessibility variables which may increase the significance of X_{11}. This would, however, result in the omission of a class of land price determining characteristics which have proven to be very important in previous studies. Unless some improvements are made with respect to elevation measurement, it appears that the use of Model 1 to measure flood damages is not feasible.

As explained above, Model 2 uses a dummy variable to differentiate between flood plain and non-flood properties. The advantage of this formulation is, of course, that the coefficient is a direct estimate of the discount in land value due to flooding. The four regression equations using Model 2 are presented in Table 2. The coefficients on the dummy variable in each equation have the expected sign and are statistically significant at the 10 percent significance level or less. Among the other explanatory variables, lot size, the average value of improvements on the block where the sample lot is located, and the distance to Sutherlin's CBD are all significant and have the expected sign in each of the equations. The overall regressions are all statistically significant at the 10 percent level. These results indicate that Model 2 is better suited for estimating residential inundation damages than Model 1.

It is necessary to choose which sample (combined or improved lots only) is to be used with Model 2. All equations (2a, 2b, 2c, and 2d) present about the same picture with respect to the independent variables.

Table 2. Results of Regression Model 2 With Two forms of the Dependent Variable and Two Subsamples.

Equation	Constant	X_1	X_2	X_3	X_5	X_7	X_9	X_{10}	X_{11}	X_{12}	R^2	N	F
Dependent Variable = Land Value Per Lot													
(2a) All Lots in Sample	133.1 (0.63)	0.04* (4.85)	-0.0005 (0.07)	-0.19* (3.66)	-0.007 (0.24)	1.15 (0.43)	0.02* (3.64)	0.10 (0.002)		-167.6* (2.91)	.533	52	6.14*
(2b) Improved Lots Only	-1.87 (0.01)	0.05* (5.57)	0.00002 (0.003)	-0.14* (2.82)	-0.003 (0.97)	3.38 (1.25)	0.02* (3.03)	-13.92 (0.33)		-198.4* (3.52)	.570	48	6.46*
Dependent Variable = Land Value Per Square Foot													
(2c) All Lots in Sample	0.07 (1.56)		-0.000003 (1.44)	-0.00004* (3.61)	-0.000002 (0.25)	-0.0004 (0.63)	0.000005* (3.37)	0.0012 (0.12)		-0.042* (3.41)	.428	52	4.71*
(2d) Improved Lots Only	0.05 (1.14)		-0.0000008 (0.58)	-0.00002* (2.88)	-0.000009 (1.54)	0.001* (2.28)	0.000003* (2.64)	-0.005 (0.68)		-0.05* (5.47)	.536	48	6.60*

* indicates statistical significance at the $(1 - \alpha) = .90$ level or greater.

Lot size, distance to the Sutherlin CBD, average value of residential improvements on the block, and the flood plain indicator are significant and have the expected sign. Variables which are insignificant in each of the equations include the value of improvements on the parcel and the distance to the southern access. Distance to I-5 is significant and positively signed in equation 2d, which is inconsistent with a priori expectations, and insignificant in the other three equations in Table 2.

From the discussion above it is obvious that the results with respect to the independent variables are about the same in all four equations. Hence, it is not possible to select a sample based on conformance with a priori expectations. The choice of the sample to use will, therefore, be made on statistical criteria. The coefficient of multiple determination and the F statistic are larger when only improved data are included in the sample and, therefore, the improved lots only sample will be used.

One form of the dependent variable will be selected to accomplish the objective of this study. The choice of equations is between 2b and 2d. In equation 2b in Table 2 the dependent variable is in terms of value per lot while equation 2d expresses the dependent variable in terms of value per square foot. Notice that X_{12} is highly statistically significant in each equation although more so in equation 2d. However, more of the variation in land value is explained by equation 2b than 2d. The overall significance of equations 2d (as shown by the F statistic) is only slightly higher than equation 2b.

The choice between the two forms of the dependent variable is not clear cut. Because of the importance of X_{12} to this study one may favor 2d slightly over 2b based on the t statistic. However, both coefficients are significant at levels greater than 99 percent. This type of confidence seems high enough for a study of this nature. Also, there are some empirical advantages associated with equation 2b. In benefit estimation one needs to know only the number of lots in the flood plain (instead of the total square feet) when value per lot is used as the dependent variable. For this reason the value per lot will be chosen for use as the dependent variable.

The best form of Model 2 must be chosen for benefit estimation. That is, one is forced to make a tradeoff between economic theory and statistical significance. Four of the independent variables presented in equation 2b in Table 2 are not significant. However, they were included in the model because economic theory indicates that each of them as well as the other dependent variables should be significant price determining characteristics. Statistical problems of correlation between independent variables make it difficult to sort out the effects that these variables have individually on land prices. The search for the best functional form of Model 2 to be discussed below is with consideration of the dictates of economic theory and the need for meaningful statistical interpretation of the independent variables, especially the flood plain dummy variable.

A compromise between economic theory and statistical significance is reached in this study through application of a two-stage process. First, several equations with differences in the number and expression of the independent variables are screened according to pre-determined statistical criteria. The second stage consists of choosing one equation from the equations which have passed the statistical screening stage. The choice of the final equation will be made solely on the basis of economic considerations.

The statistical screening is accomplished through the use of a forward selection process. In this process independent variables are added one at a time to the equation in successive steps. The independent variable with the highest absolute t value is included in the equation on the first step. At each successive step all independent variables which are not in the equation are evaluated for inclusion. Again the independent variable with the highest t value will be added at each step. This process continues until all independent variables are in the equation.[2]

After the forward selection process is complete and all independent variables are in the equation three criteria are applied to select the "best" equation or equations. These criteria are to select an equation for evaluation in stage two if it has the lowest estimated error variance (MSE), the lowest Cp statistic, or the highest R^2 adjusted for degrees of freedom in the forward selection run. These criteria attempt to determine when an "optimum" set of independent variables are included in an equation. The set is optimum in the sense that to add additional variables would impair the statistical significance of the equation. The three criteria will not always select the same equation.[3] This means that one forward

---

[2] The reader interested in a more detailed discussion of the forward selection process is referred to Draper and Smith [3], p. 169-171.

[3] For a discussion of each of the three criteria see Neter and Wasserman [8], p. 376-381. The reader may be less familiar with the Cp criteria as it is relatively new. The Cp statistic is calculated as follows:

$$Cp = \frac{SSE_p}{MSE(full)} - (n - 2p)$$

where p equals the number of estimated coefficients in the regression equation. Thus, SSE_p denotes the sum of errors squared when p-1 independent variables are in the equation. MSE(full) is the estimated variance when all independent variables to be considered are in the equation. N equals the number of observations. It is assumed that all of the independent variables have been carefully chosen so that an unbiased estimate of the true variance is obtained. If Cp is equal to or less than p then the set of independent variables is considered to provide unbiased results in the sense that the E(Yi) according to the fitted model is approxi- (continued) selection run may result in more than one equation being selected for consideration in stage two where economic appeal will be the sole criteria for determining the best equation.

Three sets of independent variables were used in separate forward selection runs. These sets are: all variables except the squared accessibility variables, all variables except the first order accessibility variables, and all variables. This division allows the effectiveness of first order accessibility variables to be compared with second order accessibility variables. The high correlation between these two groups might preclude such a comparison if only the all independent variables set were explored.

The above procedures need to be applied to equations estimated from 1962 and 1978 data in order to allow the observation of flood control benefits. This is done in the following two sections.

Selecting the Functional Form of the 1962 Equation

The results of the three forward selection runs using 1962 data and the three sets of independent variables discussed above are presented in Tables A-1, A-2, and A-3 in Appendix A. Based on these results three equations emerge for consideration in stage two. These equations are shown in Table 3. Multicollinearity is not thought to present problems in interpretation of the coefficients. The correlation coefficients are listed in Table A-4 of Appendix A.

A priori expectations with respect to lot size are realized in each of the three equations in Table 3. Land value is directly and signifi-

---

3/ cont'd. mately equal to the E(Yi) according to the true regression relation. The Cp criteria selects the set of independent variables which have the smallest calculated Cp statistic.

Table 3. Equations Selected by One or More Criteria in Three Forward Selection Runs, 1962 Data.[a]

Equation Number	Constant	X_1	X_3	X_4	X_5	X_7	X_9	X_{12}	R^2
1	47.10	0.05		-0.00006			0.02	-164	.578
	(0.48)	(6.17)		(5.20)			(3.70)	(3.96)	
2	190.1	0.05	-0.17				0.02	-155	.546
	(1.83)	(5.73)	(4.71)				(3.27)	(3.64)	
3	-28.65	0.05	-0.14		-0.03	3.64	0.02	-200	.569
	(0.15)	(5.82)	(2.98)		(1.09)	(1.46)	(3.15)	(3.80)	

[a] Criteria for selection was to choose each equation with either the lowest Cp, the lowest MSE, or the highest R^2 adjusted for degrees of freedom. Three sets of independent variables were used in three forward selection runs. Absolute values of t statistics are in parentheses.

cantly related to lot size at high levels of significance. Evidently theoretical expectations concerning the value of improvements located on a parcel hold true as this variable was not significant enough to pass the statistical screening of stage one. This would imply that appraisal data is reasonably accurate in separating from the total value of the residential package the value associated with the characteristics of the improvements on the lot and the value of the characteristics associated with the lot.

Distance to the Sutherlin CBD is significant and correctly signed in equations two and three and does not appear in equation one. The second order distance to the Sutherlin CBD variable does appear in equation one and not in equations two and three. Its coefficient is very significant but has the opposite expected sign. This may be because the range of data does not include parcels far enough from the CBD of Sutherlin to observe the expected decrease in the marginal value of this accessibility characteristic.

The distance to the southern access appears only in equation three. The coefficient has the expected sign indicating that parcels located closer to the southern access are valued higher than parcels located farther away, ceteris paribus. However, the coefficient does not appear to be a very important price determining characteristic.

Equation three is the only equation which contains distance to I-5 as an explanatory variable. The variable is significant beyond the 15 percent level of significance. However, the sign attached to the coefficient is opposite to a priori expectations. This result is likely to be caused by other characteristics which the I-5 variable is picking up. It may be that parcels located farther from I-5 have a better view, for example.

The average value of residential improvements on a block has the expected sign and is significant in all three equations. The variable indicating a flood plain location is also significant at high levels of significance and signed as expected in all three equations. This result indicates that there is a discount associated with a flood plain location. The estimated flood plain discount ranges from $155 per lot in equation two to $200 per lot in equation three.

Since the coefficient on the flood plain indicator variable is very significant in all three equations, selection of any one equation cannot be based on this variable. However, after discussion of the other independent variables, equation three seems to be most consistent with a priori economic considerations. Equation three is, therefore, selected as the 1962 equation to be used in estimating the benefits from flood control. This equation will be referred to as the 1962 equation throughout the remaining sections of this study. Plots of residuals and a brief discussion of the validity of the 1962 equation are presented in Appendix B.

Selecting the Functional Form of the 1978 Equation

The same process in which the 1962 model was selected in the last section was repeated to arrive at a "best" 1978 model. The same three sets of independent variables were included in a forward selection process. The forward selection runs are presented in Tables A-5, A-6, and A-7 in Appendix A. The results of the runs described in these appendix tables are presented in Table 4.

The absolute value of the t statistic associated with the coefficient on X_{12} is less than one in all of the equations. The insignificance of X_{12} may be because flood control has equalized land values between flood

Table 4. Equations Selected by One or More Criteria in Three Forward Selection Runs, 1978 Data.[a]

Equation Number	Constant Terms	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{12}	R^2	F
1	4793	0.44		-1.00		1.00		-64.85				-30.24	.829	59.2
	(8.61)	(14.50)		(6.46)		(8.99)		(7.85)				(0.17)		
2	3118	0.43			-.0003		.0002		-0.49			-171.0	.777	42.5
	(7.72)	(12.41)			(4.89)		(6.88)		(6.54)			(0.86)		
3	2287	0.44		-0.92		1.03			-0.45			42.6	.829	59.3
	(7.02)	(14.56)		(6.16)		(9.04)			(7.86)			(0.23)		
4	2079	0.44		-1.00		1.03			-0.46			22.61	.833	49.8
	(5.56)	(14.57)		(6.11)		(9.05)			(7.94)			(0.12)		

[a] Criteria for selection was to choose each equation with either the lowest Cp, the lowest MSE, or the highest R^2 adjusted for degrees of freedom. Three sets of independent variables were used in three forward selection runs. Absolute value of t statistics are in parentheses.

plain and non-flood plain properties. However, if this was the case it seems unlikely that X_{12} would have been entered in the model during the forward selection process. The flood plain dummy variable may become insignificant as other independent variables are added to the equation because of high correlation between X_{12} and the entering variables. Reviewing the correlation between X_{12} and the other independent variables indicates that multicollinearity may be a problem with X_7 and X_8. Each have correlation coefficients with X_{12} of close to 0.5 (see Table A-8).[4/] Before concluding that flood control has eliminated the threat of damages to residential units in the Sutherlin area the forward selection process must be run again without X_7 and X_8 in the set of variables from which to choose.

As a result of rerunning the forward selection process of stage one excluding X_7 and X_8, five equations were selected for consideration as the 1978 equation. These equations are listed in Table 5. The forward selection runs from which these equations were selected are described in Tables A-9, A-10, and A-11 in Appendix A. These five equations are substantially different, with respect to X_{12} from the four equations selected when X_7 and X_8 were not restricted from use in the model. The coefficient on X_{12} in each of the equations listed in Table 5 are negatively signed and very significant. These findings indicate that flood plain parcels are valued less than non-flood plain properties. This value differential only becomes apparent when X_7 and X_8 and the associated problems of multicollinearity with X_{12}, are removed from the model.

[4/] In the intermediate steps of each forward selection a large decrease in the absolute t value of X_{12} was observed when either X_7 or X_8 entered the model

Table 5. Equation Selected by One or More Criteria in The Forward Selection Runs Using 1978 Data and Excluding X_7 and X_8 from Consideration.

Equation Number	Constant Term	X_1	X_2	X_3	X_4	X_5	X_6	X_9	X_{10}	X_{12}	R^2	F
1	382.0 (0.97)	0.45 (10.78)				0.270 (3.29)			507.4 (2.75)	-878 (4.87)	.669	31.3
2	640.9 (1.48)	0.45 (10.95)		-0.28 (1.42)		0.36 (3.50)			409 (2.09)	-944 (5.11)	.680	25.9
3	805.2 (2.38)	0.44 (10.41)					0.00005 (2.70)		477 (2.52)	-953 (5.21)	.652	29.0
4	3886 (2.8)	0.44 (12.11)		-3.95 (4.45)	0.001 (4.32)	0.36 (0.85)	0.00003 (0.28)	-0.05 (1.51)	194 (1.08)	-865 (4.33)	.776	25.2
5	3927 (2.87)	0.44 (12.04)	-0.01 (1.01)	-3.96 (4.46)	0.001 (4.32)	0.35 (0.82)	0.00003 (0.32)	-0.04 (1.21)	183 (1.01)	-855 (4.28)	.780	22.5

a/ Criteria for selection was to choose each equation with either the lowest Cp, the lowest MSE, or the highest R^2 adjusted for degrees of freedom. Three sets of independent variables were used in three forward selection runs. Absolute value of t statistics are in parentheses.

The selection of one of the five equations listed in Table 5 for use in flood control benefit estimation will be made on the basis of how the results of each equation correspond to a priori expectations based on economic theory. Lot size is very significant and positively signed in each of the equations. This is consistent with the 1962 results and economic theory.

The value of improvements on the parcel only appears in equation 5 where it is insignificant. This result is what is expected and, as pointed out in the discussion of the 1962 results, indicates that appraisal data accurately discerns the value associated with characteristics of the improvements and the value associated with characteristics of the land. An equation which does not contain this variable is better than one that does from an economic point of view.

The distance to the Sutherlin CBD is present in equations 2, 3, and 5. The coefficient has the expected sign in each of the equations but is much more significant in equations 4 and 5. This finding is consistent with the theory which indicates distance to centers of economic activity should be significant price determining characteristics and with the 1962 results.

The second order distance to the Sutherlin CBD variable is present in equations 4 and 5. In each of these equations the coefficient is signed as expected and highly significant. This is consistent with a priori expectation that the influence of the CBD of Sutherlin on land prices would approach zero as one moves further and further from the CBD. These results are not consistent with the results of 1962 where the variable in question was found to be negatively signed in one equation and did not appear in the other two. However, the 1978 results in equation 4 and 5 are consistent with a priori expectation and, therefore, are considered to be better with respect to X_4 than are equations 1, 2, and 3.

The distance to the southern access point is in every equation with the exception of number 3. However, the coefficient has the opposite expected sign and is highly significant in equation 1 and 2. It may be that this variable is picking up the effects of other characteristics not in the equation. When addition variables are added to those contained in equations 1 and 2 the significance of X_5 declines (see equation 4 and 5). This indicates that part of the effects of other characteristics have been removed from X_5. Recall that X_5 is negatively signed in the 1962 equation. However, the coefficient is not very significant. This implies that, although the sign of X_5 is different in the 1962 equation and equation 4 and 5 of 1978, the results of the two years with respect to X_5 are not very different.

The second order distance to the Sutherlin access variable (X_6) is in equations 3, 4, and 5. As expected, the coefficient is positively signed but is significant in equation three only. On the basis of X_6, equation 3 is preferred over the others.

The average value of improvements on the block (X_9) appears in equation 4 and 5. The sign is opposite of a priori expectations. However, the coefficient does not appear to be very significant. These results are in conflict with the 1962 results in which X_9 is correctly signed and significant.

The only variable other than lot size which is present in all five equations is the non-residential land use indicator variable. The coefficient ranges from being significant in the first three equations to insignificant in equations 4 and 5 at standard levels of probability.

A priori expectations are for a negative sign on the coefficient of X_{10}. However, all of the coefficients on X_{10} in Table 5 are greater than zero. These findings are different from the 1962 results where X_{10} was insignificant and did not appear in any of the equations.

The 1978 results with respect to X_{10} indicate that location next to a non-residential land use may have an enhancing effect on land values. Recall from Chapter III that X_{10} was used to depict a location next to one or more of the following land uses: commercial, industrial, schools, parks, and churches. The expected sign on X_{10} was negative because the congestion associated with each of the above land uses was thought to detract from the quality of the site. However, each of these land uses has positive attributes associated with it as well. For example, schools, parks and churches usually have wide open areas of green space while commercial and industrial land uses provide easy access to employment and shopping needs. It seems, at least in the 1978 equations, that the positive attributes of these types of land uses outweigh any negative effects.

Based on the above discussion equation 4 in Table 5 seems to be most consistent with a priori expectations. This equation will, therefore, be used along with equation 3 in Table 2 to estimate the residential benefits from flood control on Sutherlin Creek. The selection of any equation other than 4 would not effect the results of the benefit estimation to any large degree. This is so because the flood plain coefficient is nearly the same in all five equations ranging from -953.0 in equation 3 to -855.0 in equation 5. Plots of the residuals and a brief discussion of the appropriateness of the 1978 equation are presented in Appendix B.

Benefit Estimation

Using Model 2 for the residential inundation benefit estimation, the coefficient on the flood plain dummy variable may be considered as the discount in price associated with flood plain parcels. Since the discount reflects the expected damages from flooding two discounts were estimated: one to reflect the expected damages without the project and the other with the project.

These coefficients need to be made comparable by eliminating the effect of changes in the general level of land prices in Sutherlin between 1962 and 1978. An index was developed for this purpose representing the ratio of the data set mean land value in 1978 over the mean land value in 1962. This ratio is equal to 8.0 which implies that the 1962 discount of $200 per lot is equivalent to a discount of $1,600 per lot in 1978 dollars.

The damages per lot without the project are expected to be $1,600 while remaining damages are equal to $865 per lot (in 1978 dollars). The benefits from flood control are equal to $735 per lot. This measure of benefits per lot must now be applied to the number of lots located in the flood plain.

The flood risk discount was applied to 235 parcels representing the total number of residential lots within the flood plain before the project (1967). Note that this also includes several houses scattered throughout the flood plain which were not located in blocks or subdivisions and hence were not part of the regression analyses. The flood risk discount (i.e., inundation damages and inconveniences), is assumed to be the same as in residential blocks. Table 6 gives a breakdown of this total by lots located within the boundaries of residential blocks and subdivisions and those residential units situated within the flood plain but outside of

Table 6. Location and Number of All Residential Units Located Within the Boundaries of the Sutherlin Creek Flood Plain in 1967.

Residential Areas	Improved Lots	Unimproved Lots	Total Lots
Sherwood Addition	2	0	2
East Sutherlin Terrace	5	2	7
Mardonna Homesites	5	5	10
Area South of H.S. Track	8	1	9
Maple J. Abby	2	7	9
Grace Subdivision	2	5	7
Gourley's Addition	8	7	15
Block 22	12	4	16
36	7	3	10
21	14	7	21
24	7	2	9
31	2	2	4
38	7	1	8
20	1	2	3
39	4	0	4
40	4	3	7
41	2	1	3
33	11	4	15
34	17	9	26
35	6	2	8
Outliers	42	0	42
Total residential units within the flood plain in 1967			235

clearly defined residential areas. The latter types are referred to as outliers.

Total inundation reduction benefits (in 1978 dollars) are $172,725 which equals the reduced damages per lot because of flood control ($735) multiplied by the number of residential units within the Sutherlin Creek flood plain (235).

The total inundation reduction benefits reflect two things. First, there is the expected value of the reduction in annual inundation damages. In addition, this measure includes the value of reduction in inconveniences associated with a flood plain location. Both of these types of annual benefits are discounted by residents of the flood plain. This process of discounting the annual benefits yield the figure of $172,725 which represents total benefits.

To express total benefits in annual terms the process of discounting annual benefits must be reversed. To accomplish this it is necessary to know how long the flow of benefits are expected to last and the rate at which annual benefits are discounted. It is assumed that residents expect these benefits to be forthcoming during the expected life of the project which is 100 years.

Assuming that the mortgage rate is the appropriate discount rate makes it possible to estimate annual benefits. The mortgage rate for 1978 was just under 9.7 percent. Based on a 100 year flow of benefits, a 9.7 percent amortization rate, and total benefits of $172,725
tial inundation reduction benefits are equal to $15,275.

CHAPTER V

SUMMARY AND CONCLUSIONS

Results and Implications

The objective of this study was to use implicit price equations to estimate residential benefits from flood control on Sutherlin Creek, Oregon. Two variables were considered for measuring the flood hazard associated with parcels in the flood plain. These variables were elevation above mean sea level and an indicator variable depicting a flood plain location. It was found that the indicator variable was an effective measure of the flood hazard while the elevation variable was not.

To arrive at an estimate of flood control benefits it was necessary to estimate two cross-sectional implicit price equations. The expected damages without flood control were estimated using appraisal data for 1962. The estimated equation is:

$$\begin{array}{lllllll} Y = -28.65 & + \; 0.05X_1 & - \; 0.14X_3 & - \; 0.03X_5 & + \; 3.64X_7 & + \; 0.02X_9 & - \; 200X_{12} \\ & (5.82) & (2.98) & (1.09) & (1.46) & (3.15) & (3.80) \end{array}$$

where

Y = appraised lot value in 1962

X_1 = lot size in square feet

X_3 = distance in feet from sample lot to the Sutherlin CBD

X_5 = distance in feet from sample lot to the southern access

X_7 = distance in feet from sample lot to Interstate-5

X_9 = average appraised value of improvements on the block

X_{12} = indicator variable set equal to one if sample lot is located in the flood plain and zero otherwise.

The 1962 results indicate that flood plain parcels are valued \$200 less than non-flood plain parcels, ceteris paribus. The difference of \$200 can be interpreted as the discounted flow of annual flood damages associated with residential parcels within the Sutherlin Creek flood plain. Damages are likely to consist of physical costs (e.g., replacement cost of flood damaged carpet) and psychological costs (e.g., the mental anguish of living with the threat of flooding).

The year 1978 was selected to analyze the impact of flood control on the discounted flow of annual damages. The estimated implicit price equation for 1978 is:

$$\begin{aligned} Y = 3886 &+ \underset{(12.11)}{0.44X_1} - \underset{(4.45)}{3.95X_3} + \underset{(4.32)}{0.001X_4} + \underset{(0.85)}{0.36X_5} + \underset{(0.28)}{0.00003X_6} \\ &- \underset{(1.51)}{0.05X_9} + \underset{(1.08)}{1.94X_{10}} - \underset{(4.33)}{865X_{12}} \end{aligned}$$

where

X_1, X_3, X_5, X_9, and X_{12} are as previously defined.

Y = appraised lot value in 1978

X_4 = the squared distance in feet from sample lot to the Sutherlin CBD

X_6 = the squared distance in feet from sample lot to the southern access

X_{10} = indicator variable set equal to one when sample parcel is located next to a nonresidential land use.

The results of the 1978 land price equation indicate that flood control has not reduced discounted damages from flooding to zero. Remaining damages with flood control are estimated to be \$865 per lot. However, after inflating the 1962 flood plain discount by a factor of eight to reflect the increase in the value of residential land in the Sutherlin area between 1962 and 1978 it becomes apparent that damages have been reduced. The 1962 estimated damages per lot of \$200 is equivalent to \$1,600 per lot in 1978 terms. It is, therefore, estimated that flood control on Sutherlin Creek has resulted in a reduction in damages equal to \$735 per lot in 1978 terms.

Total benefits from flood control of \$172,725 were estimated by applying the \$735 to the 235 residential parcels in the Sutherlin Creek flood plain. This must be interpreted as a minimum estimate of benefits. The reason for this restriction is that the \$735 represents the amount consumers would have to pay for a flood-free parcel (as reflected in appraisal data) and not the amount that consumers would be willing to pay.

Total benefits were annualized by assuming a rate of discount and a time period over which the flow of benefits were expected to be forthcoming. The average mortgage rate for 1978 was approximately 9.7 percent and was used as the discount rate. A 100 year flow of annual benefits, which is equal to the expected life of the project was used also to estimate annual benefits. Based on these assumptions annual benefits from flood control to residential property are equal to \$15,275.

Conclusions

It has been assumed in this study that annual flood damages are capitalized into the value of land. The hypothesis of this paper was that these damages could be estimated with implicit price equations.

The regression results indicate that residential lots exposed to the flood hazards of Sutherlin Creek are valued less than lots not exposed to the threat of flooding. This apparent value differential may be because the flood plain indicator variable, as hypothesized, is reflecting the discounted value of expected damages associated with a flood plain location. However, this variable may also be reflecting the effects of other land price determining characteristics not in the implicit price equations.

Selection of the land price determining characteristics to be included in the implicit price model was accomplished with a great deal of thought. Four categories of characteristics were considered with the objective of choosing those characteristics thought to be important land price determining variables within the Sutherlin area. Because of the care taken to include all important characteristics in the land price model the author cannot think of any potentially significant variables that may have been excluded from the model. It is, therefore, concluded that the coefficients on the flood plain indicator variables are reflecting the discounted value of expected flood damages with and without flood control.

Suggestions for Future Research

It has been demonstrated that flood control benefits can be estimated using land value regression models. However, several questions regarding the relationship between this technique and the damage-frequency technique come to mind. For example, it would be of interest to know if flood control benefits estimated with implicit price models are likely to be higher, lower, or approximately equal to estimates derived from standard damage-frequency techniques. Also, the relationship between the discharge-frequency characteristics of a watershed and the comparability of the two estimates would be of interest. That is, would frequent flooding enhance the comparability of estimates derived from either of the two methods? If so, how frequent must flooding events be to yield similar benefit estimates between the damage-frequency method and the implicit price method? Research into the relationship between these two methods of estimating flood control benefits could prove useful in understanding these relationships.

LITERATURE CITED

[1] Brigham, Eugene F. "The Determinants of Residential Land Values". Land Economics. November 1965.

[2] Damianos, Demetrios and Leonard A. Shabman. "Land Prices in Flood Hazard Areas: Applying Methods of Land Value Analysis". Bulletin 95, Virginia Water Resources Research Center. Virginia Polytechnic Institute and State University, April 1976.

[3] Draper, Norman and Harny Smith. Applied Regression Analysis. John Wiley and Sons, Inc., New York, London, and Sydney, 1966.

[4] Freeman, Myrick A. III. "On Estimating Air Pollution Control Benefits From Land Value Studies". Journal of Environmental Economics and Management 1. 1974.

[5] Jaksch, John A. and Herbert H. Stoevener. "Effects of Air Pollution on Residential Property Values in Toledo, Oregon". Agricultural Experiment Station, Oregon State University, Special Report 304, September 1970.

[6] Kates, Robert William. Hazard and Choice Perception in Flood Plain Management. University of Chicago. Department of Geography. Research Paper No. 78. 1962.

[7] Knetsch, Jack L. "Influence of Reservoirs Projects on Land Values". Journal of Farm Economics. Vol. 46, No. 1. February 1964.

[8] Neter, John and William Wasserman. Applied Linear Statistical Models. Richard D. Irwin, Inc. Homewood, Illinois. 1974.

[9] Schefter, John Edwin. "The Demand for and Supply of the Characteristics of a New Residence and the Residential Location Decision". Unpublished Ph.D. dissertation. Department of Agricultural and Resource Economics, Oregon State University, July 1977.

[10) Soil Conservation Service. Sutherlin Watershed Project Work Plan. Douglas County, Oregon, 1963.

[11] Stull, William J. "Community Environmental Zoning and the Market Value of Single Family Homes". The Journal of Law and Economics, Vol. 18, No. 2, October 1975.

APPENDIX A

Table A-1. Criteria For Selecting the 1962 Model with X_1, X_2, X_3, X_5, X_7, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a/]	Variable Added	P[b/]	Cp	MSE	R^2	R^2adj
1	X_1	2	24.07	21,672	.249	.233
2	X_3	3	20.16	20,231	.314	.284
3	X_{12}	4	11.4	17,114	.433	.394
4*	X_9	5	3.15	14,016	.546	.504
5	X_7	6	4.24	14,032	.556	.503
6*	X_5	7	5.11	13,970	.569	.505
7	X_{10}	8	7.00	14,278	.570	.494
8	X_2	9	9.00	14,645	.570	.482

[a/] * indicates that the model meets one or more of the three criteria.

[b/] P equals the number of estimated coefficients in the regression at each step.

Conclusions: The Cp criterion favors stopping with step 4. The MSE and R^2adjusted criteria favor stopping at step 6.

Table A-2. Criteria for Selecting the 1962 Model with X_1, X_2, X_4, X_6, X_8, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a/]	Variable Added	P[b/]	Cp	MSE	R^2	R^2adj
1	X_1	2	31.3	21,672	.249	.232
2	X_4	3	26.27	20,087	.319	.289
3	X_{12}	4	15.85	16,805	.443	.405
4*	X_9	5	4.35	13,038	.5777	.539
5	X_{10}	6	6.00	13,239	.581	.531

[a/] Forward selection was terminated by the computer at step 5 because X_2, X_6, and X_8 have entering $|t|$ values of 0.10 or less. * indicates the model meets one or more of the three criteria.

[b/] P equals the number of estimated coefficients in the regression at each step.

Conclusion: All three criteria favor stopping at step four.

Table A-3. Criteria for Selection of the 1962 Model with X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a]	Variable Added	P[b]	Cp	MSE	R^2	R^2adj
1	X_1	2	28.12	21,672	.24918	.233
2	X_4	3	23.4	20,087	.319	.289
3	X_{12}	4	13.49	16,805	.443	.405
4*	X_9	5	2.56	13,038	.5777	.538
5	X_3	6	4.08	13,190	.583	.533
6	X_7	7	5.69	13,382	.587	.526
7	X_{10}	8	7.44	13,629	.589	.518
8	X_6	9	9.00	13,823	.594	.511

[a] Forward selection was terminated by the computer at step 8 because X_2, X_5, and X_8 have entering $|t|$ values of 0.10 or less. * indicates that the model meets one or more of the three criteria.

[b] P equals the number of estimated coefficients in the regression at each step.

Conclusion: All three criteria indicate the best model is reached at step four.

Table A-4. Correlation Coefficients 1962 Sample.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{12}
Y	.499	.233	-.304	-.298	-.202	-.166	.034	.014	.154	.083	.040
X_1		.098	-.100	-.067	-.249	-.194	-.038	-.030	-.045	.122	.433
X_2			-.113	-.127	.000	.043	.118	.110	.095	.015	-.067
X_3				.980	.618	.652	-.151	-.093	.399	-.338	-.441
X_4					.586	.611	-.276	-.216	.434	-.297	-.434
X_5						.969	.374	.419	.361	-.243	-.429
X_6							.424	.475	.479	-.218	-.336
X_7								.997	.110	-.180	.347
X_8									.140	-.190	.344
X_9										-.190	-.109
X_{10}											.118

Table A-5. Criteria for Selection of 1978 Model with X_1, X_2, X_3, X_5, X_7, X_8, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a]	Variable Added	P[b]	Cp	MSE	R^2	R^2adj
1	X_1	2	130.8	697,695	.446	.438
2	X_{12}	3	81.3	520,119	.594	.581
3	X_5	4	71.1	483,055	.628	.611
4	X_7	5	43.8	380,290	.712	.694
5*	X_3	6	4.81	229,429	.829	.815
6	X_9	7	5.8	229,437	.832	.815
7	X_{10}	8	7.3	231,212	.833	.814
8	X_2	9	9.0	233,981	.834	.811

[a] * indicates that the model meets one or more of the three criteria.

[b] P equals the number of estimated coefficients in the regression at each step.

Conclusion: All three criteria dictate stopping at step 5.

Table A-6. Criteria for Selecting the 1978 Model with X_1, X_2, X_4, X_6, X_8, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a/]	Variable Added	P[b/]	Cp	MSE	R^2	R^2adj
1	X_1	2	83.1	697,695	.446	.438
2	X_{12}	3	46.2	520,119	.594	.581
3	X_6	4	42.2	498,774	.616	.598
4	X_8	5	25.0	410,346	.689	.669
5*	X_4	6	3.9	299,488	.777	.759
6	X_{10}	7	5.4	301,908	.779	.757
7	X_9	8	7.23	306,317	.779	.753
8	X_2	9	9.0	310,366	.780	.750

[a/] * indicates that the model meets one or more of the three criteria.

[b/] P equals the number of estimated coefficients in the regression at each step.

Conclusion: All three criteria dictate stopping at step 6.

Table A-7. Criteria for Selection, 1978 Model with X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a/]	Variable Added	P[b/]	Cp	MSE	R^2	R^2adj
1	X_1	2	127.1	697,695	.446	.438
2	X_{12}	3	78.5	520,119	.594	.581
3	X_5	4	68.6	483,054	.628	.611
4	X_8	5	38.0	365,694	.723	.705
5*	X_3	6	3.6	229,156	.829	.815
6*	X_9	7	4.39	228,153	.832	.816
7	X_{10}	8	5.75	229,464	.835	.815
8	X_6	9	7.38	231,872	.836	.813
9	X_2	10	9.14	234,935	.936	.811
10	X_4	11	11.0	238,545	.837	.808

[a/] Forward selection was termined at step 10 because X_7 had an entering $|t|$ value of 0.10 or less. * indicates that the model meets one or more of the three criteria.

[b/] P equals the number of estimated coefficients in the regression at each step.

Conclusion: Cp favors stopping at step 5. MSE and adjusted R^2 favor stopping with step 6.

Table A-8. Correlation Coefficients 1978 Sample.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{12}
Y	.688	-.099	.028	.092	.114	.086	-.234	-.221	.035	.121	-.105
X_1		-.013	-.120	-.101	-.191	-.121	.064	.066	.001	.036	.375
X_2			.160	.169	.170	.219	.211	.227	.395	-.139	.097
X_3				.982	.675	.685	-.016	.050	.531	-.465	-.313
X_4					.633	.635	-.121	-.050	.576	0.423	-.314
X_5						.972	.453	.501	.431	-.324	-.219
X_6							.526	.577	.518	-.318	-.087
X_7								.997	.189	-.167	.499
X_8									.234	-.189	.496
X_9										-.276	.017
X_{10}											.096

Table A-9. Criteria for Selection of the 1978 Model with X_1, X_2, X_3, X_5, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a]	Variable Added	P[b]	Cp	MSE	R^2	R^2adj
1	X_1	2	40.9	697,695	.446	.438
2	X_{12}	3	15.2	520,119	.594	.581
3	X_5	4	10.7	483,055	.628	.611
4*	X_{10}	5	5.117	437,393	.669	.647
5*	X_3	6	5.123	430,291	.680	.653
6	X_2	7	6.4	432,592	.683	.651
7	X_9	8	8.0	436,567	.686	.648

[a] * indicates that the model satisfies at least one of the three criteria.

[b] P equals the number of estimated coefficients in the regression equation.

Conclusions: The Cp criteria favors stopping at step 4. The MSE and adjusted R^2 criteria favor stopping at step 5.

Table A-10. Criteria for Selection of the 1978 Model with X_1, X_2, X_4, X_6, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a/]	Variable Added	P[b/]	Cp	MSE	R^2	R^2adj
1	X_1	2	35.43	697,694	.446	.438
2	X_{12}	3	11.2	520,119	.594	.581
3	X_6	4	9.2	498,774	.616	.598
4*	X_{10}	5	4.85	459,564	.652	.630
5	X_2	6	6.	460,713	.657	.629

[a/] * indicates that the model satisfies at least one of the three criteria. Process stops at step 6 because $|t|$ of X_4 and X_9 is less than 0.10.

[b/] P equals the number of estimated coefficients in the regression equation.

Conclusion: All three criteria favor stopping at step four.

Table A-11. Criteria for Selection of the 1978 Model with X_1, X_2, X_3, X_4, X_5, X_6, X_9, X_{10}, X_{12} Included in a Forward Selection Process.

Step[a]	Variable Added	P[b]	Cp	MSE	R^2	R^2adj
1	X_1	2	80.7	697,695	.446	.438
2	X_{12}	3	44.5	520,119	.594	.581
3	X_5	4	37.4	483,055	.628	.611
4	X_{10}	5	28.9	437,393	.669	.647
5	X_6	6	24.2	409,983	.695	.670
6	X_9	7	24.9	409,680	.700	.670
7	X_3	8	25.7	410,487	.704	.669
8*	X_4	9	9.03	315,757	.776	.745
9*	X_2	10	10.	315,589	.780	.746

[a] * indicates that the model satisfies at least one of the three criteria.

[b] P equals the number of estimated coefficients in the regression equation.

Conclusions: The Cp criterion dictates stopping with step 8. The MSE and R^2 adjusted criteria dictates stopping with step 9.

APPENDIX B

APPENDIX B

VALIDITY OF THE 1962 AND 1978 MODELS

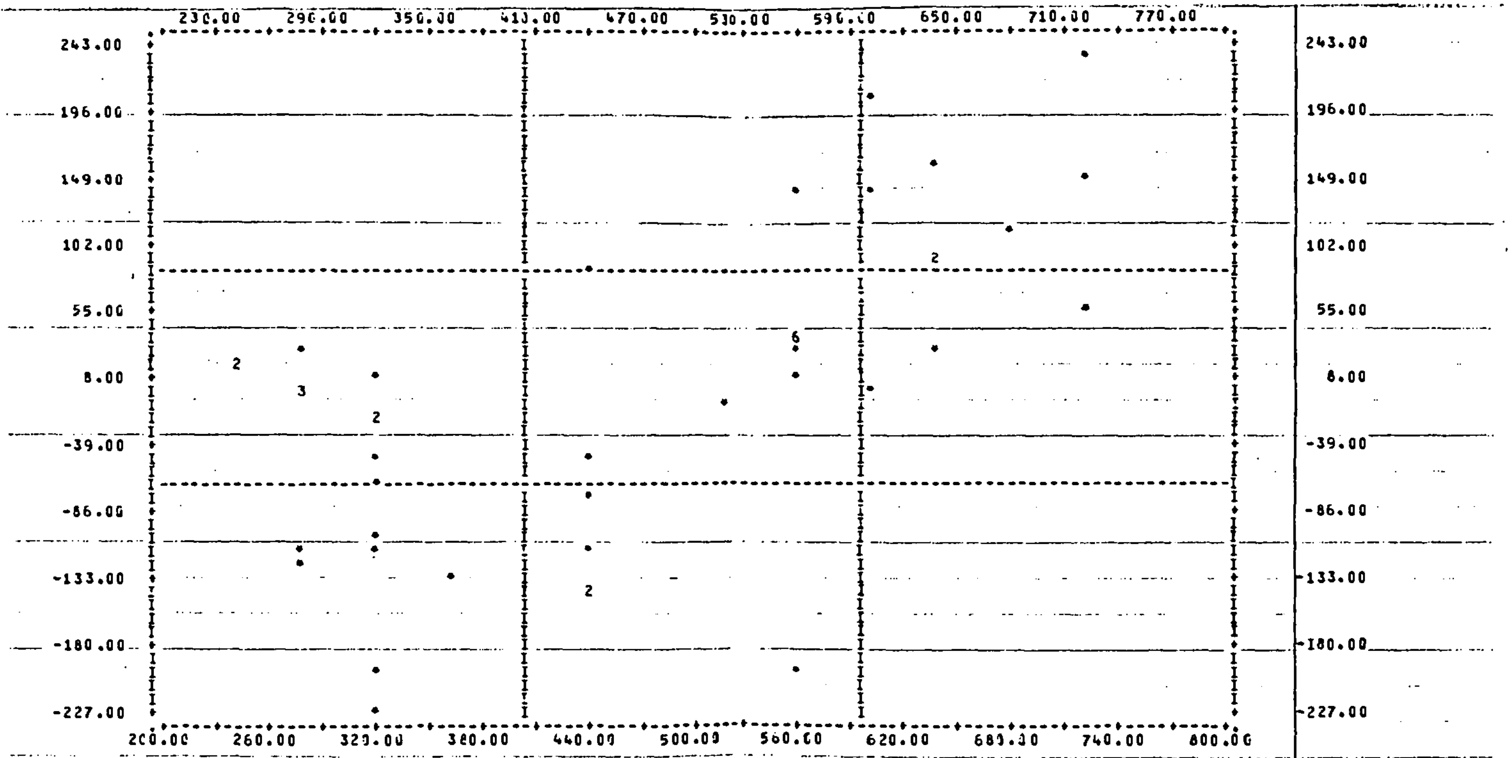
The Validity of the 1962 Model

In Figu a plot of the residuals against the dependent variable, land value p, is presented. This plot has several implications. First, it apthat the variance of the error term is approximately equal over ages of the dependent variable. That is, heteroskedasticity is not ered to be a problem.

A distufeature about Figure B-1 is the apparent positive trend in the error. This trend is really seen by an examination of the statistics abottom of Figure B-1. The correlation coefficient between the als and lot value is about 0.6. A line fitted by the ordinary leares technique has a positive slope of 0.43 and is highly signi

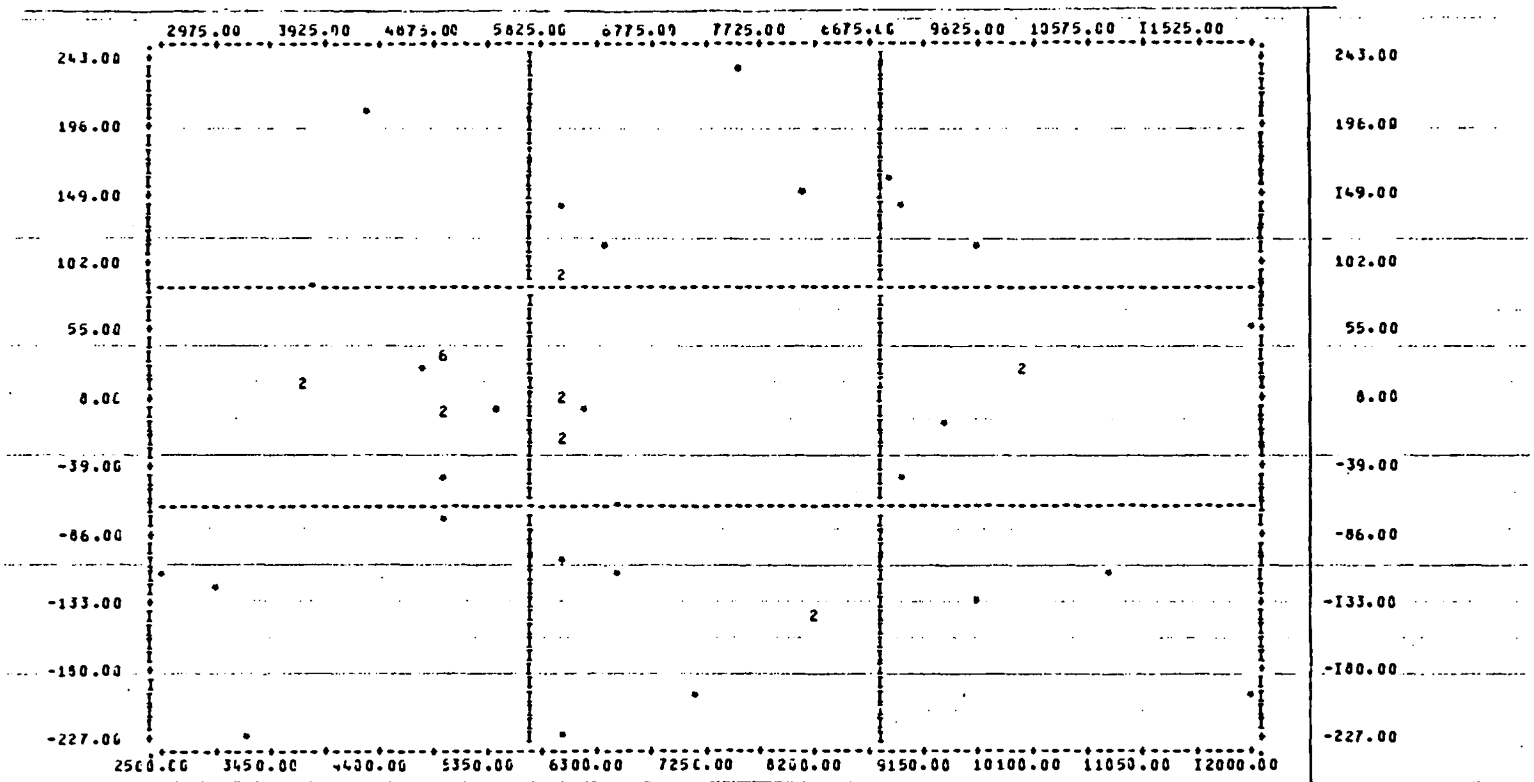
The resplot of Figure B-1 indicates that an important variable has possibly excluded from the 1962 equation. It may be possible to include a vain the equation which would eliminate the positive correlation betwe residuals and lot price. However, all independent variables (ansformations of these variables) which were thought to be significan price determining characteristics in the Sutherlin area with thetion of characteristics for which data was not available, were sd for consideration in the land value model. The comprehensive a prilection of potentially important independent variables makes it diffto identify what variable or variables may have been left out of tel.

Figures -3, and B-4 are residual plots against lot size, the



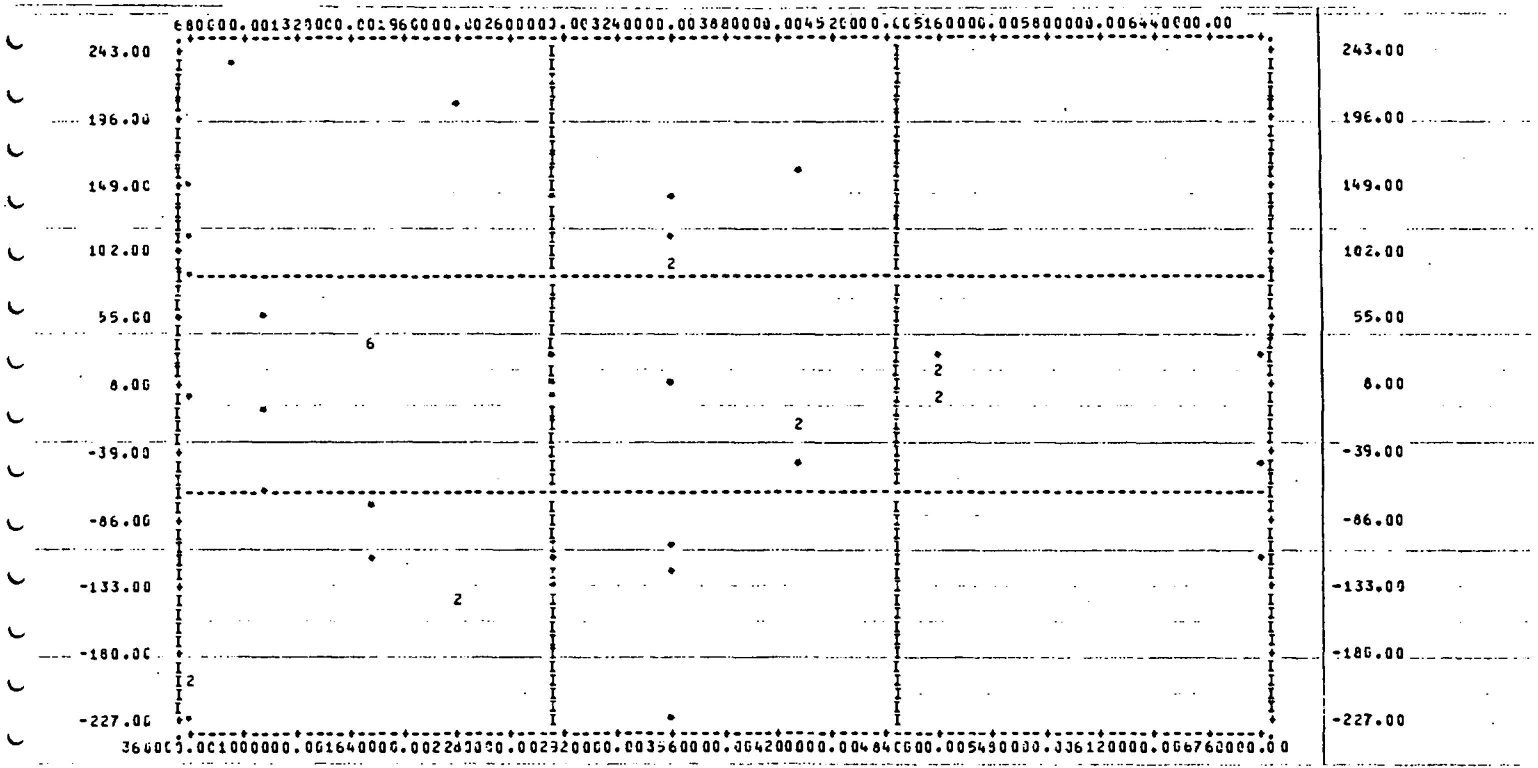
```
STATISTICS..
  CORRELATION (R)-         .61299      R SQUARED      -        .37575      SIGNIFICANCE R -        .00001
  STD ERR OF EST -       85.03136      INTERCEPT (A)  -    -198.70398      STD ERROR OF A -      40.99925
  SIGNIFICANCE A -         .00001      SLOPE (B)      -        .43264      STD ERROR OF B -        .08407
  SIGNIFICANCE B -         .00001
```

Figure B-1. Plot of Residuals (vertical axis) Against 1962 Land Value Per Lot.



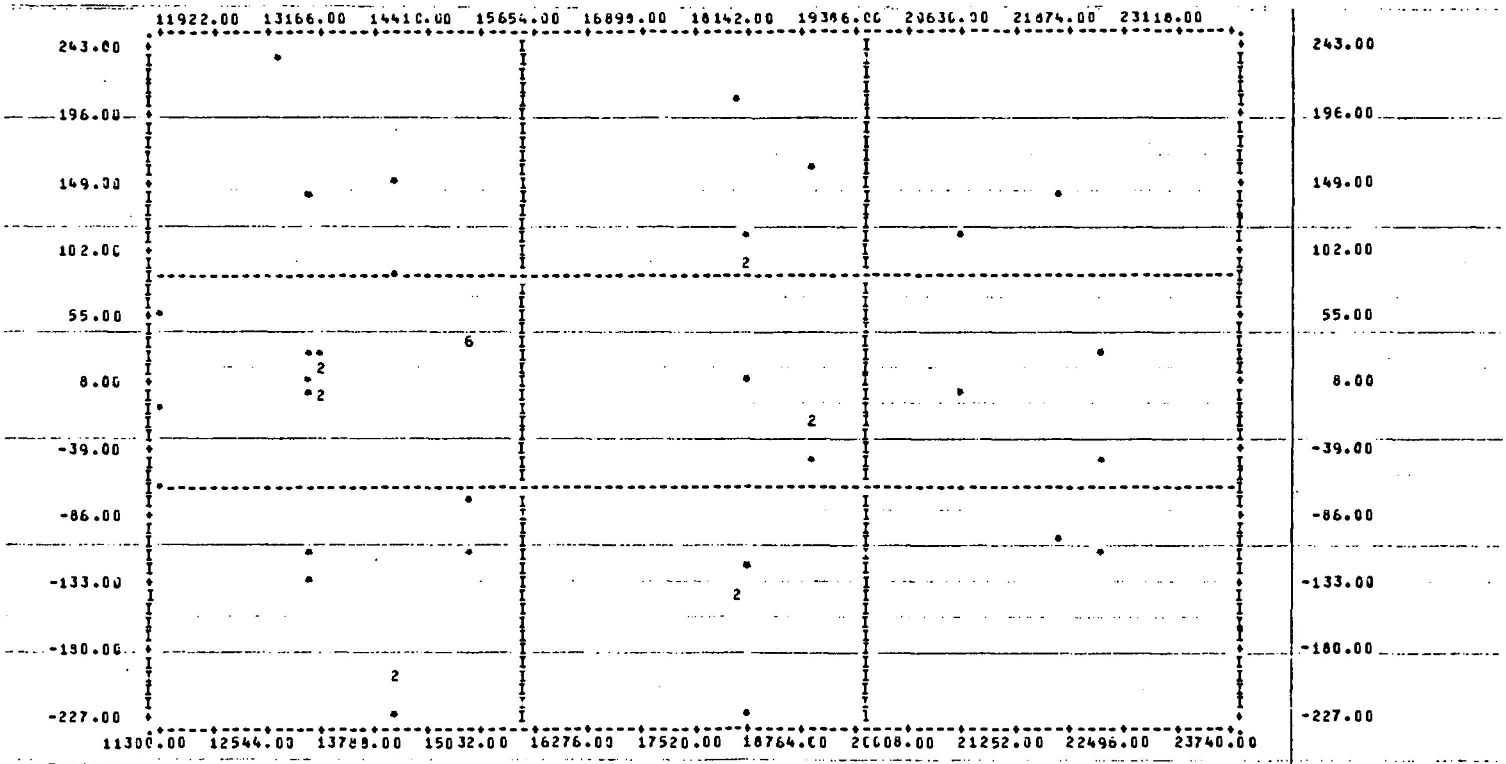
```
STATISTICS..

  CORRELATION (R)-          .00000     R SQUARED       -        .00000     SIGNIFICANCE R -         .50000
  STD ERR OF EST -       111.58714     INTERCEPT (A)   -   -.46230E-05     STD ERROR OF A -       48.51900
  SIGNIFICANCE A -          .50000     SLOPE (B)       -    .59587E-09     STD ERROR OF B -         .00701
  SIGNIFICANCE B -          .50000
```

Figure B-2. Plot of Residuals (vertical axis) Against Lot Size.



```
STATISTICS..

  CORRELATION (R)-        -.01711      R SQUARED       -          .00029      SIGNIFICANCE R -         .45406
  STD ERR OF EST -      111.57031      INTERCEPT (A)   -         2.75973      STD ERROR OF A -       28.71861
  SIGNIFICANCE A -         .46193      SLOPE (B)       -     -.10619E-05      STD ERROR OF B -     .91494E-05
  SIGNIFICANCE B -         .45406
```

Figure B-3. Plot of Residuals (vertical axis) Against Squared Distance to the Sutherlin CBD.



```
STATISTICS..

  CORRELATION (R)-          .00000     R SQUARED       -        .00000      SIGNIFICANCE R -          .50000
  STD ERR OF EST -       111.56714     INTERCEPT (A)   -   -.32534E-05      STD ERROR OF A -        81.90642
  SIGNIFICANCE A -          .50000     SLOPE (B)       -    .15719E-09      STD ERROR OF B -          .00500
  SIGNIFICANCE B -          .50000
```

Figure B-4. Plot of Residuals (vertical axis) Against the Average Value of Improvements on the Block.

distance to Sutherlin's CBD squared, and the average value of improvements on the block, respectively.

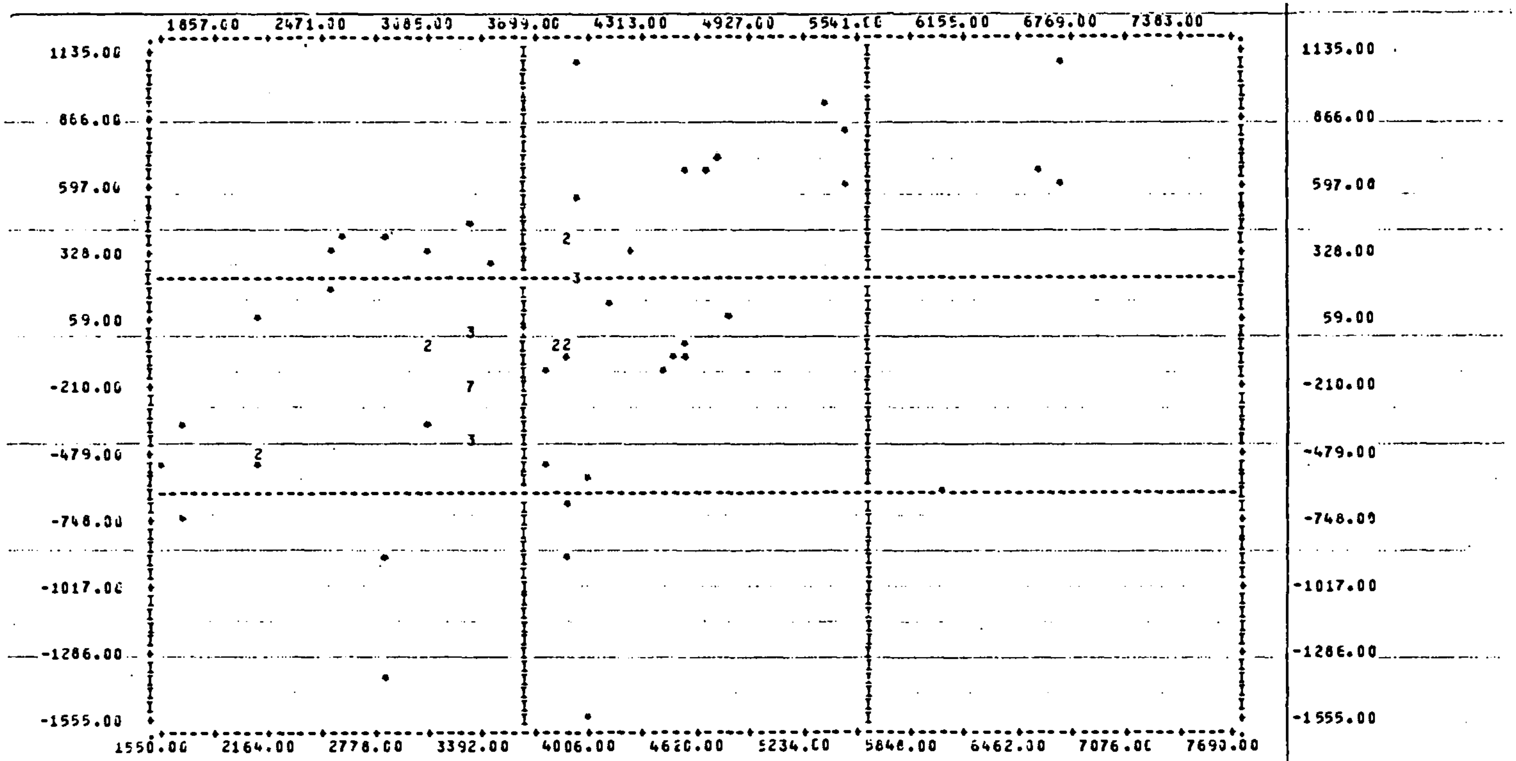
In each of these plots the variance of the residuals seem to approximately constant over all values of the independent variables. There is also no relationship between the residuals and the independent variables in the 1962 equation.

The Validity of the 1978 Model

A plot of residuals against the response variable (land value per lot in 1978) is shown in Figure B-5. This plot suggests a positive relationship between the two variables as do the statistics listed at the bottom of Figure B-5. This plot is similar to the plot in Figure B-1 for the 1962 model. It seems that an important independent variable has been omitted from the 1978 equation. However, for reasons mentioned in the discussion of validity of the 1962 model, finding the independent variable which should have been included in the land value model would be difficult.

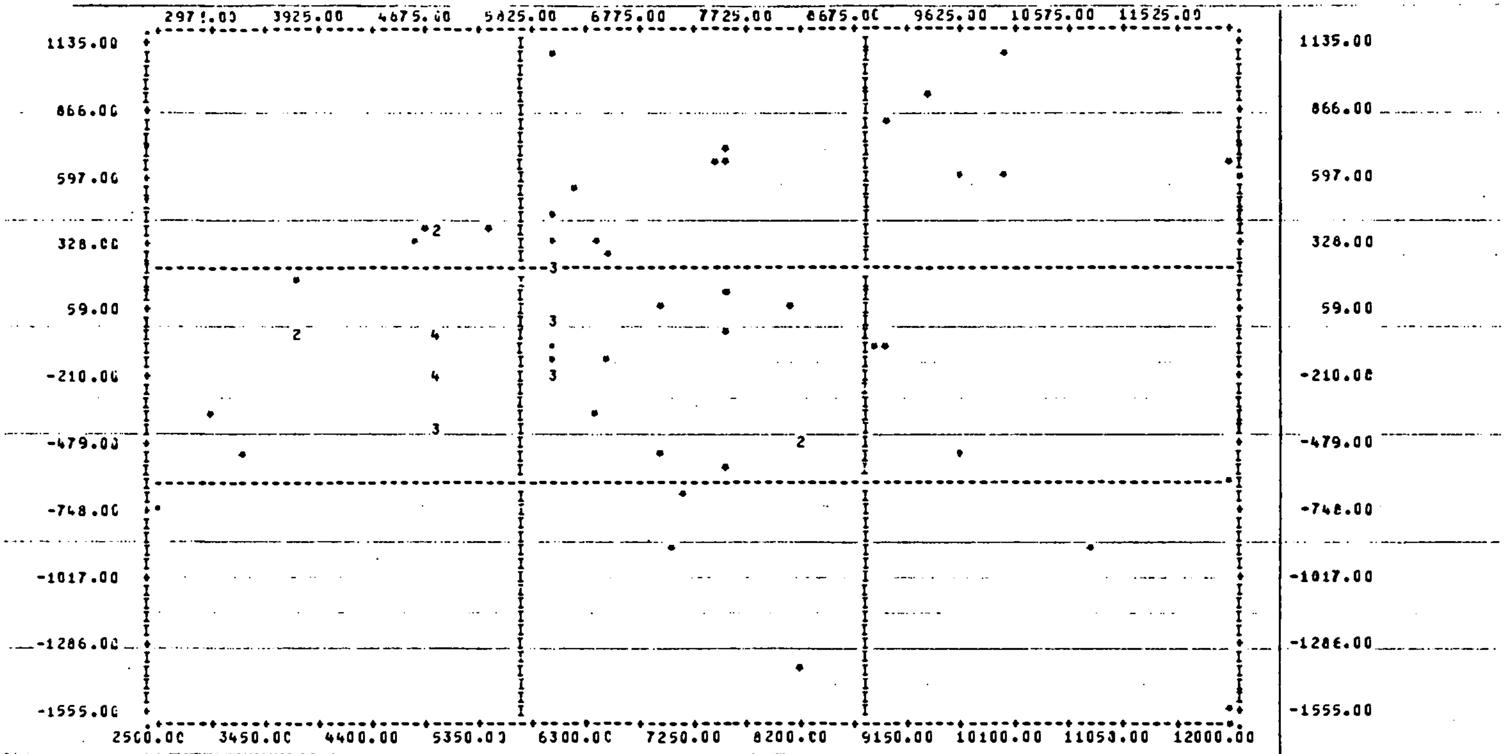
The assumption of homoskedasticity does not appear to be violated. The residuals in the first horizontal section of Figure B-5 are spread over approximately two vertical sections. About the same spread is observed in the second and third horizontal section.

Plots of residuals against the independent variables in the 1978 equation are presented in Figure B-6 through B-11. The only plot which seems abnormal is the plot against lot size in Figure B-6. The variance of the residuals seems to be increasing with the size of lot. The implication of this plot is that the variance of residuals would also increase with land value per lot (because lot size and land value are positively correlated). However, Figure B-5 does not support this.



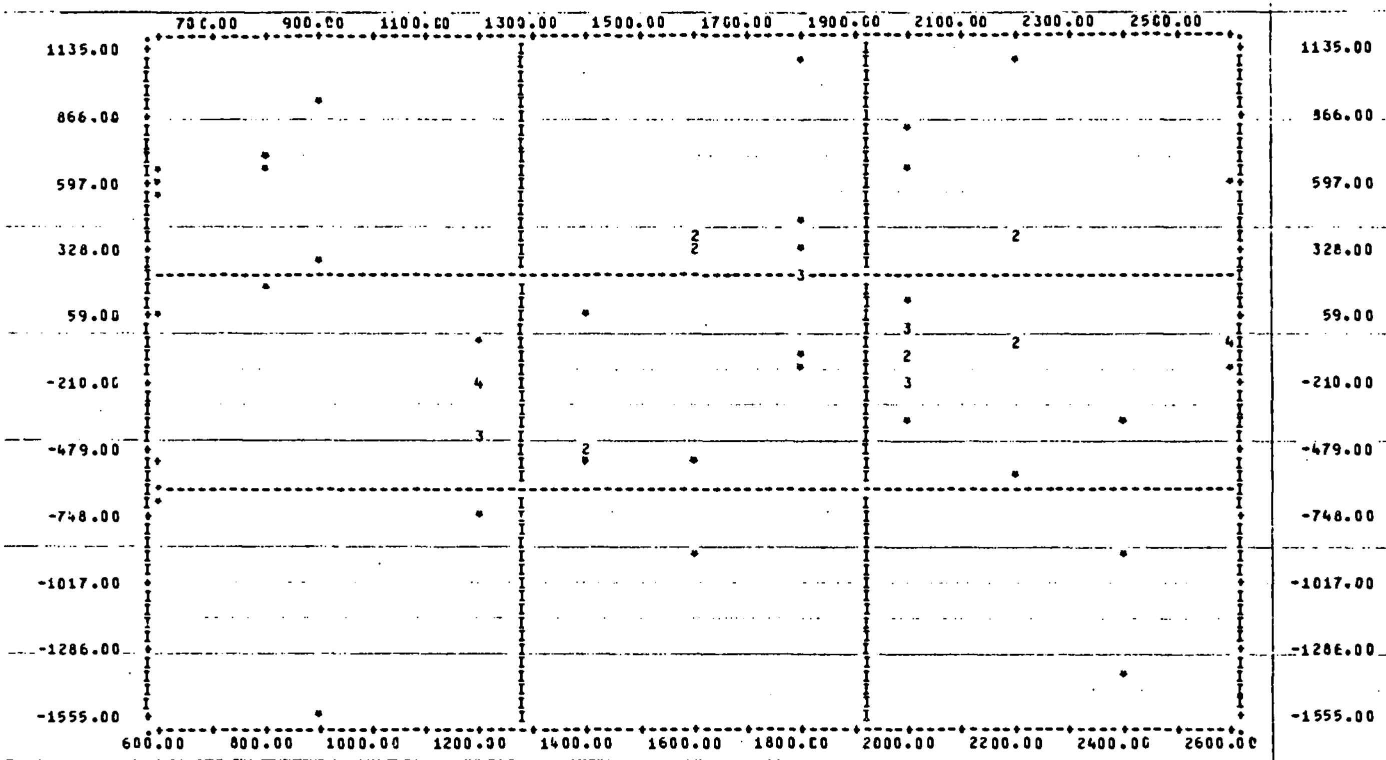
```
STATISTICS..

  CORRELATION (R)-          .47288      R SQUARED        -         .22361      SIGNIFICANCE R -         .00003
  STD ERR OF EST -       467.70575      INTERCEPT (A)    -      -831.94064      STD ERROR OF A -      200.58706
  SIGNIFICANCE A -          .00005      SLOPE (B)        -         .22361      STD ERROR OF B -         .05168
  SIGNIFICANCE B -          .00003
```

Figure B-5. Plot of Residuals (vertical axis) Against 1978 Land Value Per Lot.



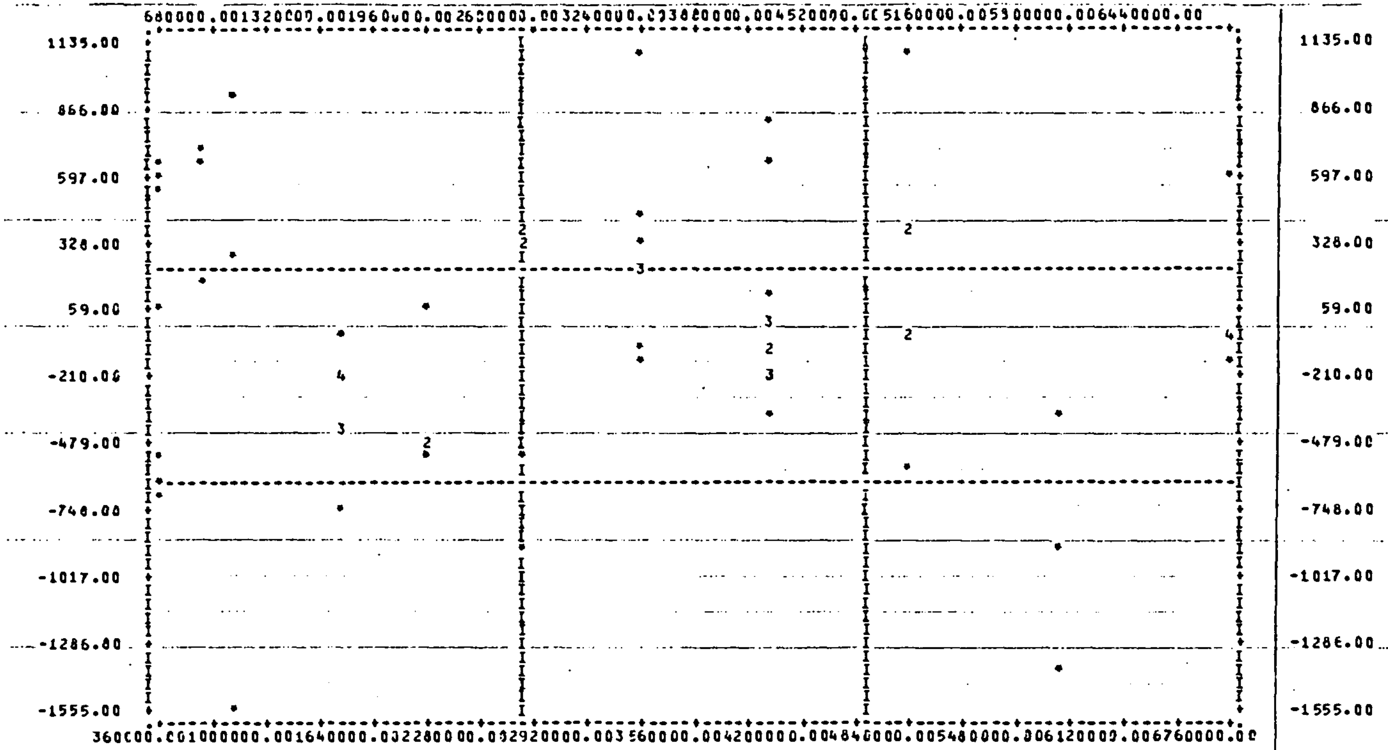
```
STATISTICS..

  CORRELATION (R)-         -.00000      R SQUARED        -          .00000      SIGNIFICANCE R -       .50000
  STD ERR OF EST -       530.80323      INTERCEPT (A)    -          .00006      STD ERROR OF A -    213.39213
  SIGNIFICANCE A -          .50000      SLOPE (B)        -     -.11074E-07      STD ERROR OF B -       .03078
  SIGNIFICANCE B -          .50000
```

Figure B-6. Plot of Residuals (vertical axis) Against Lot Size.



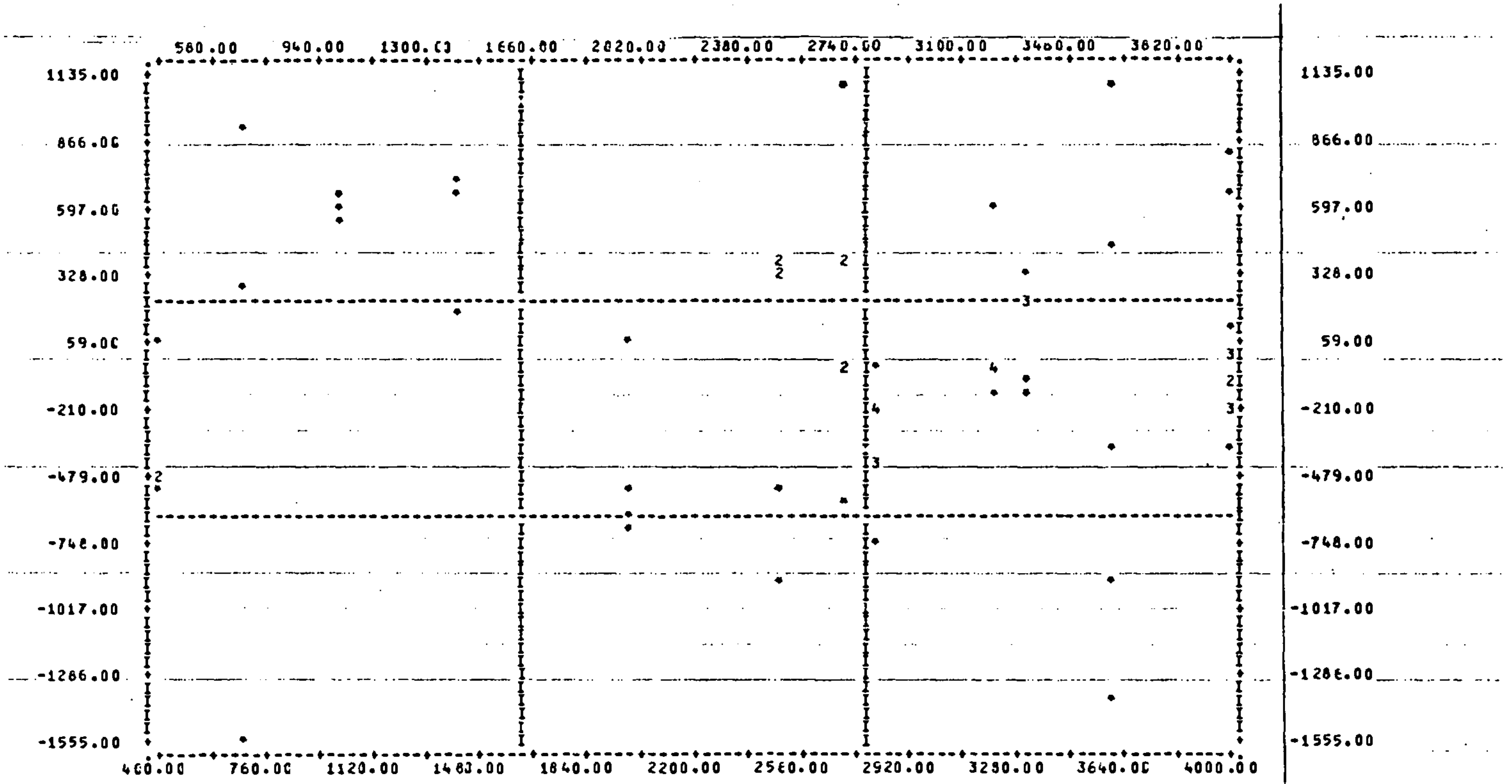
```
STATISTICS..
  CORRELATION (R)-          .00000      R SQUARED       -        .00000      SIGNIFICANCE R -        .50000
  STD ERR OF EST -       530.80328      INTERCEPT (A)   -       -.00002      STD ERROR OF A -     188.13631
  SIGNIFICANCE A -          .50000      SLOPE (B)       -    .63903E-08      STD ERROR OF B -        .10786
  SIGNIFICANCE B -          .50000
```

Figure B-7. Plot of Residuals (vertical axis) Against Distance to Sutherlin's CBD.



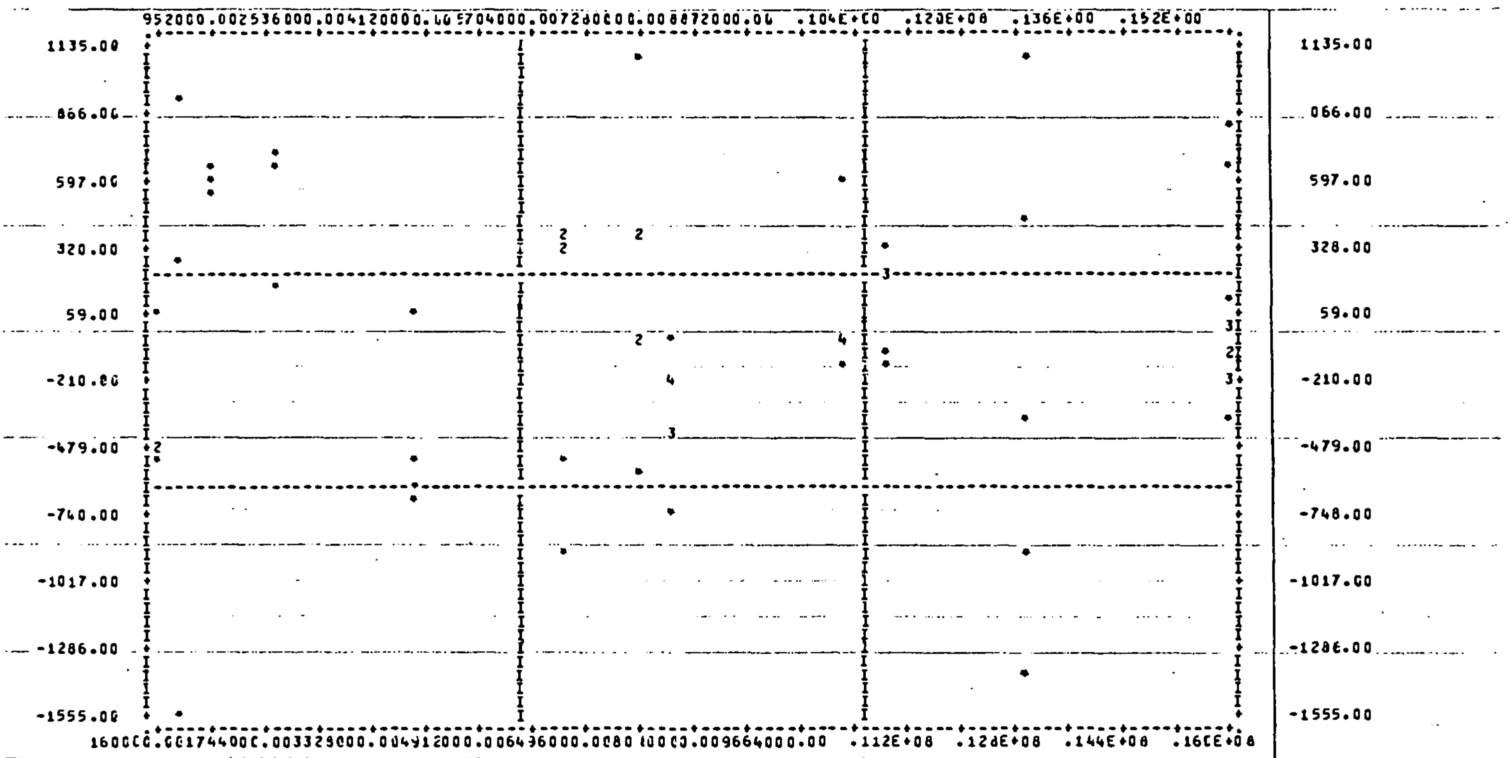
```
STATISTICS..
  CORRELATION (R)-          .00000     R SQUARED       -         .00000     SIGNIFICANCE R -          .50000
  STD ERR OF EST -       530.80328     INTERCEPT (A)   -       -.00002     STD ERROR OF A -       121.39494
  SIGNIFICANCE A -          .50000     SLOPE (B)       -         .00000     STD ERROR OF B -          .00003
  SIGNIFICANCE B -          .50000
```

Figure B-8. Plot of Residuals (vertical axis) Against Squared Distance to Sutherlin's CBD.



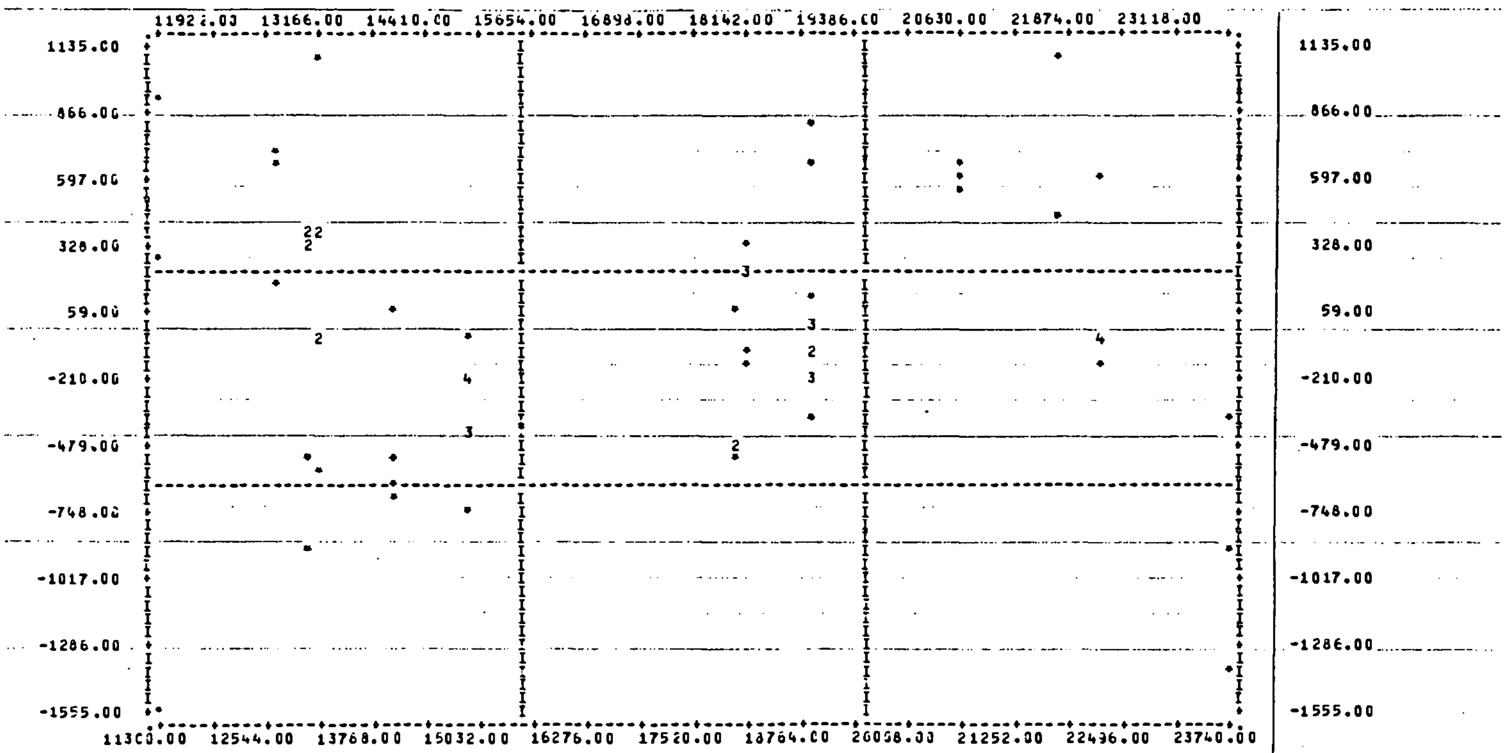
```
STATISTICS..
  CORRELATION (R)-          .00000     R SQUARED       -          .00000     SIGNIFICANCE R -          .50000
  STD ERR OF EST -       530.80328     INTERCEPT (A)   -         -.00004     STD ERROR OF A -       175.05318
  SIGNIFICANCE A -          .50000     SLOPE (B)       -     .10625E-07     STD ERROR OF B -          .06042
  SIGNIFICANCE B -          .50000
```

Figure B-9. Plot of Residuals (vertical axis) Against Distance to Southern Access.



```
STATISTICS..

  CORRELATION (R)-          .00300       R SQUARED       -          .00000       SIGNIFICANCE R -         .50300
  STD ERR OF EST -       530.89328       INTERCEPT (A)   -        -.00003       STD ERROR OF A -      125.53695
  SIGNIFICANCE A -          .50000       SLOPE (B)       -          .00000       STD ERROR OF B -         .00001
  SIGNIFICANCE B -          .50000
```

Figure B-10. Plot of Residuals (vertical axis) Against Squared Distance to Southern Access.



```
STATISTICS..

  CORRELATION (R)-          .00000      R SQUARED        -         .00000      SIGNIFICANCE R -        .50000
  STD ERR OF EST -       530.80328      INTERCEPT (A)    -        -.00008      STD ERROR OF A -     316.37616
  SIGNIFICANCE A -          .50000      SLOPE (B)        -     .38798E-08      STD ERROR OF B -        .01824
  SIGNIFICANCE B -          .50000
```

Figure B-11. Plot of Residuals (vertical axis) Against the Average Value of Improvements on the Block.