The Effect of the Washington Metro on Urban Property Values

Steven R. Lerman
David Damm
Eva Lerner-Lamm
Jeffrey Young

Massachusetts Institute of Technology

JULY, 1978

FINAL REPORT

DOCUMENT IS AVAILABLE TO THE U.S. PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

Prepared for

U. S. DEPARTMENT OF TRANSPORTATION
URBAN MASS TRANSPORTATION ADMINISTRATION
Washington, DC 20590

494565

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle: "The Effect of the Washington Metro on Urban Property Values"		5. Report Date: November, 1977
		6. Performing Organization Code
7. Author(s): Steven R. Lerman, David Damm, Eva Lerner Lam, Jeffrey Young		8. Performing Organization Report No.: CTS Report No 77-18
9. Performing Organization Name and Address: Massachusetts Institute of Technology, Centre for Transportation Studies		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.: MA-11-0004
12. Sponsoring Agency Name and Address: Urban Mass Transportation Administration, U.S. Department of Transportation		13. Type of Report and Period Covered
		14. Sponsoring Agency Code

15. Supplementary Notes

Literature Review as Appendix

16. Abstract

As with other forms of urban infrastructure, public mass transit systems can alter the spatial distribution of urban property values. The magnitude of this effect is likely to be highly parcel-specific, and changes in real estate values may occur both prior to and after a transit system's construction.

This report describes a series of econometric models of real estate values estimated for parcels in Washington, D.C. over the period of the Metro system's development. Separate models are estimated for single family dwellings, multi-family structures and retail stores. Access to the transit system and the implementation schedule of Metro are both found to be significant determinants of parcel transaction prices.

The first chapter of the report is a general overview and summary of conclusions; it can be read independently of the remainder of the report.

17. Key Words	18. Distribution Statement
-Property values -Transit Impacts -Washington Metro -Transit Financing	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
		140	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Massachusetts Institute of Technology
CENTER FOR TRANSPORTATION STUDIES

CTS Report Number 77-18

November, 1977

THE EFFECT OF THE WASHINGTON METRO ON URBAN PROPERTY VALUES

by

Steven R. Lerman
David Damm
Eva Lerner-Lam
Jeffrey Young

Prepared for the
Urban Mass Transportation Administration
U.S. Department of Transportation
Washington, D.C.

METRIC CONVERSION FACTORS

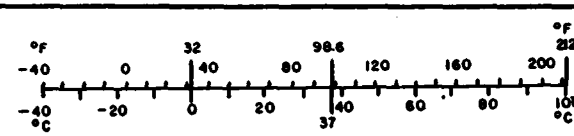
Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in ≡ 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



LIST OF TABLES

Table No	Title	Page
2.1	Station Area Characteristics	14
2.2	Characteristics of Tracts Outside Station Areas	16
2.3	Number of Years to Completion	18
2.4	Summary of Single Family Dwelling Sample	21
2.5	Summary of Multi-Family Building Sample	22
2.6	Summary of Retail Establishment Sample	23
3.1	Variables in the Models	30
3.2	Choice of Variables for Single Family Dwelling Model	38
3.3	Elasticities of Price with Respect to Distance to Station	50
3.4	Elasticities of Price with Respect to Years to Completion	51
4.1	Potomac Avenue Sample Statistics	56
4.2	Base Case Model Results for Potomac Avenue	57
4.3	Percentage Differences in Model Forms for Potomac Avenue	58
4.4	Potomac Avenue: Effect of Delay in Station Opening	59
4.5	Potomac Avenue: Effect of Relocation of Station	61
4.6	Potomac Avenue: Effect of Elimination of Station	62
4.7	Profile of Multi-Family Case Studies' Station Areas	64
4.8	Sample Characteristics for Multi-Family Building Case Studies	64
4.9	Base Case for Multi-Family Dwelling	65
4.10	Multi-Family Building Case Studies: Delay in Station Opening	66
4.11	Multi-Family Building Case Studies: Relocation of Station	67
4.12	Dupont Circle Sample Characteristics	68
4.13	Dupont Circle: Effect of Delay in Station Opening	69
4.14	Dupont Circle: Effect of Relocation of Station	70

TABLE OF CONTENTS

Section	Page
1. INTRODUCTION	
1.1 - Study Objectives	1
1.2 - The Approach of the Study	3
1.3 - Summary of Major Findings	6
1.4 - Structure of Report	8
2. DESCRIPTION OF THE DATA	9
2.1 - Parcel Data	10
2.2 - Station Areas' Characteristics	12
2.3 - Auxiliary Data	15
2.4 - Summary of Samples	19
2.5 - Evaluation of Samples	24
3. EMPIRICAL RESULTS	26
3.1 - Model Specification Process	26
3.2 - Definition of Variables	29
3.3 - Single Family Dwelling Models	32
3.4 - Multi-Family Building Models	40
3.5 - Retail Establishment Models	44
3.6 - Summary and Evaluation of Model Estimation Results	50
4. CASE STUDIES OF STATION AREAS	53
4.1 - Potomac Avenue/Single Family Base Case	55
4.2 - Potomac Avenue/Single Family Case Study	59
4.3 - Multi-Family Building Case Studies	63
4.4 - Dupont Circle/Retail Establishment Case Study	68
4.5 - Summary of Case Studies	71
5. CONCLUSIONS	73
REFERENCES	77
APPENDIX -A-	

1. INTRODUCTION

1.1 Study Objectives

It has long been recognized that the provision of public infrastructure has a profound influence on the pattern of urban development and the spatial distribution of urban real estate values. The existence of highways, sewer services and other public facilities influence the behavior of both suppliers and demanders of residential and commercial properties. The benefits of these facilities or services are, at least in theory, partially or wholly capitalized into urban property values.

As the cost of public investments in general and mass transit systems in particular has increased, there has been a growing interest in the concept of "taxing back" the publicly-induced real estate value increases to help finance public investments. In the context of mass transit, this concept has been generally labelled "value capture" policy, and has led to the formulation of a broad range of strategies for public recuperation of supposed upward property value shifts.

Value capture policies, including direct taxation, public acquisition of property prior to construction with subsequent resale or leasing, and joint public/private development are predicated on a number of basic assumptions. First, these policies presume that benefits accruing to individuals who own property that is affected by investments in public infrastructure should be taxed directly rather than taxing income or capital gains. Second, they assume that some special mechanism other than the usual property taxes are needed to recover transit-related benefits. However, most fundametally, value capture policies presume that transit systems actually have an impact on urban properties, and that such impacts are large enough to be worth the effort of recapturing. It is this last assumption that is the focus of this study.

Unlike several other recent and ongoing research projects, (e.g. Sharpe (1976) , Burkhardt (1976)), we will be concerned here neither with the exact definition of 'value capture' nor best or most appropriate approach to capturing values. Instead, our primary interest is to establish empirical evidence as to the influence of the existence and incidence of mass transit facilities on different types of urban properties' values.

The study has four goals. First, and most obvious, we want to determine if a transit system's construction has measurable effect on the values of urban parcels. Moreover, the focus of the study is particularly on the response of urban property values in anticipation of the implementation of a heavy rail transit system. Previous studies (e.g. Boyce, et al.(1972)) have examined transit-related effects on property values in a suburban market in at least a limited way, but pre-implementation impacts have not yet been analyzed. Second, we hope to generate information on the distribution and magnitude of value increments and, perhaps, value decrements. Third, unlike prior studies we sought a refined understanding of the effects that the presence of a major transit investment has on parcels of different market segments, i.e. single and multi-family residences and retail establishments. Fourth, we wish to explore how different transit system designs and implementation schedules might affect property value shifts in order to provide some basis for assessing how such political decisions influence the potential values to be "captured".

1.2 - The Approach of the Study

The first step toward understanding the relationship between land values and a transportation system's development was an evaluation of the literature bearing on prior work in this field. We found this work to be extensive but varying widely in quality. Appendix A provides a structured and detailed analysis of the state-of-the-art. On the basis of this review and analysis, we drew the following conclusions relevant to our particular research:

(1) Very few studies have pursured research in this area in an analytically rigorous fashion. Those which were particularly haphazard in their methodological decision-making (sample size, appropriate dependent and independent variables, data sources, overall analytical framework) tended to come out with "definitive" conclusions, i.e. hard numbers to justify their hypotheses. However, when considering their methodologies, the accuracy of their models and reliability of their predictions seemed highly questionable. On the other hand, those studies which took pains to make the careful methodological decisions tended to qualify their findings and seemed unable to commit themselves to any "definitive" conclusions on the basis of their empirical findings. This latter case illustrates the difficulty of this type of research and the elementary level of the state-of-the-art.

(2) Any methodology which relies heavily upon controls in the sample data (in any time or space dimension) is likely to be subject to problems since the real world has a tendency to violate the assumed control conditions. On the other hand, methodologies which utilize statistical controls within the model (e.g. multivariate regression techniques), enable the analyst to at least theoretically isolate and estimate the effect on the independent variables on the dependent variable.

(3) When using any type of regression technique, the key issue is not to omit any variables which may covary with the transit access term; by omitting such variables, one leaves the transit access variable to "pick up" all the effects, thus over-or underestimating the true effect of transit accessibility. Goodness-of-fit, as measured by statistics such as a high R^2 are of secondary importance so long as the transit access term and any omitted terms are not correlated.

(4) There seems to be no single, appropriate methodology for choosing the appropriate independent variables within a regression framework. Although several statistical techniques such as stepwise linear procedures, use of such methods creates serious problems of misspecification and "hiding" of potentially valuable information.

(5) The nuisance effects associated with transportation facilities have rarely been considered by analysts in the past. It is possible to incorporate variables which reflect the level of negative effects within a multivariate regression analysis.

The second task of the study was the development of a data base appropriate for a multivariate regression analysis of urban property values.

The recently opened Washington, D.C. Metro system was selected as a study site. For reasons discussed further in Chapter 2, the study area was restricted to Washington, D.C. proper rather than the entire metropolitan area.

A sample of data on real estate transactions in Washington from 1969 to 1976 was collected and augmented with information from the 1970 census and the 1974 Metropolitan Parcel File, a computer-based file describing every parcel in Washington. This data (along with some auxiliary information from maps, WMATA planning documents and price deflators based on Bureau of Labor Statistics publications) was organized into three samples: single family dwellings; multi-family buildings and retail establishments. For the retail establishments, further transactions in the 1966-1969 were added to the file. Each sample included data about the transaction (price, year of transaction), as well as information on the following three categories:

(1) parcel data - distance of parcel to Metro Center, floor area (for retail establishments), lot area, number of dwelling units (for multi-family buildings), zoning class;

(2) transit system-related - distance to nearest Metro station, whether the nearest station was above or below ground, whether the station had a park-ride lot, the number of years between the transaction and the nearest station's scheduled opening (at the time of transaction);

(3) neighborhood-related - characteristics of area around parcel including racial composition, income, housing quality, employment density, fraction owner-occupiers and population density.

The final data sets for single family dwellings, multi-family buildings and retail establishments consisted of 286, 771 and 353 observations respectively.

The general characteristics of these samples were explored in cross-tabulations and correlation analyses. Three parallel model-building efforts were subsequently conducted on the samples, each incorporating a wide range of independent variables to explain transaction prices (deflated to 1969 dollars). Alternative functional forms of both the dependent and independent variables were explored. In addition, a special set of so-called "Box and Cox" analyses of transformations of the dependent variable (price) were performed. This technique allows for a wide range of functional forms within a general class of model specifications and provides a statistical test for the appropriateness of the most widely used transformations.

Each of the parallel regression studies produced at least one model which was deemed acceptable based on intuitive considerations (reasonable signs and magnitudes of coefficients) as well as statistical criteria (goodness-of-fit, low standard errors on coefficient estimates, etc.). These models were then used in a series of station area case studies. In these case studies, specific stations were selected and the models were used to forecast real estate values. Predictions from the models were compared with observed transaction values of parcels near the station. This "base case" forecast provided as least a limited basis for judging the models' validity.

The base case was then altered to reflect the changes in the Metro system design. For example, the effect of closing the station was represented by determining the value of the variable "distance to nearest Metro station" when the current closest station was closed. Forecasts of transaction values were then made under the assumed revised conditions and compared with the base case. This analysis made possible at least a crude assessment of the magnitude of property value shifts in response to transit policies.

1.3 Summary of Major Findings*

The major conclusions of the study are as follows:

1) The distance of a parcel from a Metro station appears to be a determinant of the variation in values of properties in Washington, D.C.

* The purpose of this section is to present these findings which could be of interest to planners and public officials contemplating various value capture policies. The reader who desires a more complete account of our findings is directed to sections 3.6 and 4.5.

2) The estimated elasticity for single-family dwellings, multi-family buildings and retail establishments of deflated price with respect to distance from a station varies from -.06 to -.13 for single family dwellings. Comparable values for multi-family buildings and for retail properties are -.06 and -.13 respectively. Each of these figures is only an estimate and should be used with great care.*

3) As with distance, the opening data of a specific Metro station appears to have an effect on propery values. The value of a parcel tends to increase as the nearby stations opening data approaches.

4) The estimated elasticity of price with respect to the number of years to completion ranges from -.05 to -.11 for single family dwellings, to -.09 for multi-family buildings and -.15 for retail properties. As with the elasticities of price with respect to distance-to-station reported above, these figures are estimates only.*

5) It is uncertain whether factors related to nuisances possibly associated with Metro stations (e.g. noise, road congestion, air pollution) are significant determinants of the values of property.

6) Non-Metro-related factors (parcel-specific and demographic) seem to have substantial influence on all parcel types. Collectively, these variables have a far greater influence on real estate values than any of the transit-related factors.

7) The availability of parking positively influences retail property values. The elasticity of price with respect to retail floor area is an estimated 3.7 times the comparable elasticity with regard to parking lot area.

---

* See section 3.6 for standard errors of these estimates.

1.4 - Structure of Report

The remainder of this report is in four distinct sections and an appendix. Chapter 2 describes the data base, and includes a description of the original sources of information, how these sources were used, and a set of summary tabulations for each of the samples.

Chapter 3 is a description of thr modelling results, including both some of the preliminary results and the final models. The single-family dwelling, multi-family building and retail establishment models are presented in separate subsections.

Chapter 4 describes the case studies. Included in this chapter are studies of five separate transit stations. For each case study, base case results and forecasts of property value changes under different conditions are given.

The final chapter of the main body of the report is a summary of how this study relates to more general transit financing and planning issues.

The appendix is a self-contained review of the state-of-the-art in understanding the effect of transportation investments on property values. It is intended as background material for interested readers and other researchers in the field.

2. DESCRIPTION OF THE DATA

As the available literature described in Appendix A indicates, there have been significant problems associated with obtaining reliable data about urban real estate transactions. Obtaining a complete history of the transactions for any reasonable size sample of parcels was infeasible within the time and budget constraints of this study. Instead, we decided to utilize a sample of transactions for the period 1969-1976 with no attempt to follow any particular parcel's history. For the retail establishment sample, the period from 1969 to 1976 as augmented with data from 1966 through 1969. (This data was included in order to statistically test whether there were observed shifts in property values between the '66-'69 and '69-'76 periods).

The following three distinct data sets were developed for the study:

(1) owner-occupied single-family dwellings;

(2) multi-family buildings;

(3) retail establishments.

In each case, we restricted the sample to Washington, D.C., proper rather than the entire metropolitan area. Not only did this avoid the problems of differing conventions for recording real estate transactions in the various jurisdictions of the Washington region, but it also facilitated the entire data-gathering process.

The following sections describe the sources and assumptions used to build the data base used for estimating the real estate value models. Section 2.1 has a discussion of the two primary sources of information about specific parcels of land. In Section 2.2 we present the method used to build descriptors of each parcel's neighborhood. Section 2.3, is a brief discussion of all the auxiliary data sources, including housing prices deflators, access to transit and transit construction for the Washington Metro System.

The following section, 2.4, is a tabulation of the major characteristics of each of the final samples used for model estimation. It includes a summary of the sample means, standard deviations and ranges. In Section 2.5, some of the potential deficiencies of the final data bases are discussed along with a general evaluation of the usefulness of the data.

2.1 Parcel Data

The principal source of parcel-data was the Metropolitan Parcel File (MPF) which contained a number of valuable pieces of data on individual parcel in Washington, D.C. Since the MPF included virtually all parcels in the District, it provided the universe from which we were able to sample and locate our observations, *i.e.*, transacted parcels. In addition to having exact premise addresses, it also contained a number of useful identifiers, including 1970 U.S. Census block and tract number, Transportation Planning Board Zone, as well as a MPF-specific sequence number. The MPF also had a number of potentially interesting data describing each parcel. Unfortunately we were able only to make use of a few of these as either data was missing for a given parcel or it was clearly unreliable.

Our second source of data was directories of virtually all real estate transactions in Washington. These directories were rented from R.S. Lusk and Sons, Inc., and provided a rich source of transaction data. These directories contained extensive information on recorded transfers of all property in Washington, D. C. from 1969 to 1976. In addition to the amount of the transaction for a specific parel, the date of sale, mortgage interest rate, name of new owner, local block and lot descriptions, assessment data as well as land use type was given.

Transacted parcels were isolated by sampling from the 1974 Washington, D.C. Metropolitan Parcel File (MPF), and checking each address for transactions using the Lusk listings for 1969 to 1976*. Once a match had been made, the following data was recorded for the observation:

1) sequence number - assigned to each parcel in the MPF, it was used to identify the parcel as well as to match the LUSK data to the MPF data.

2) parcel's address (from MPF) - recorded by street's number, name and quadrant.

3) transaction price (from Lusk) - at time of transaction

4) data of transaction (from Lusk) - recorded by month and year

5) parcel characteristics - including floor area (for retail establishments only), lot area, structure condition, zoning of parcel.

---

*Parcel files from 1970, 1972, 1973 and 1974 were provided by the Washington Council of Governments.

2.2 Station Areas' Characteristics

For each Metro station in the District of Columbia, a surrounding station area was defined. The boundaries of these station areas were determined by the following criteria:

1) Station areas extend approximately 1/2 to 1 mile from the central station site;
2) Station areas were formed from complete census tracts, if possible. (This was done for ease of data collection, allowing the use of census tract statistics.)
3) Station areas were formed from portions of census tracts if condition 2 was infeasible. (The division of census tracts was determined by major physical structures such as rivers, major highways, or local arterials.)
4) In areas where stations were located in close proximity to each other (less than 1 mile apart), the boundaries of the station areas were approximately the equidistant line between the stations.

The characteristics of the station areas were used as estimates of the general conditions in the surrounding neighborhood. In all cases, the station area statistics were calculated using 1970 Census tract and block data. Definitions for station area variables are listed below.

1) % Renter = Rented Dwellings in Station Area/Housing Units in Station Area X 100
2) % Owner = Owned Dwellings in Station Area/Housing Units in Station Area X 100
3) % Non-White = Non-Whites in Station Area/Population in Station Area X 100

4) $\% \text{ Substandard} = \frac{\text{Housing Units lacking Plumbing and Other Facilities}}{\text{Housing Units in Station Area}} \times 100$

5) $$\text{Mean Income} = \frac{\sum_{i=1}^{n} (\text{Population in Census Tract } i) \times (\text{Mean Income in Tract } i)}{\text{Total Population in Station Area}}$$

where: n = the number of census tracts in station area

6) Distance to Metro Center = Distance from station in a given station area along the Metro right-of-way to the Metro Center station.

In many cases, a particular parcel in the sample was located outside of a defined station area. The neighborhood socio-economic data used for these parcels was that of the Census tract in which the parcel lay. Metro-related data was determined by the assignment of a Metro station to the transacted parcel, using the following criteria:

1) the closest Metro station by straight line distance, or
2) if a number of stations are equal distance from the parcel by straight line distance, the closest station was determined by street distance.

Table 2.1 is a summary of the station area characteristics. Along with the six station area variables defined above, the following are given:

1) SS - an index equal 1 if the station is above ground, 0 if otherwise.
2) PR - an index equal to 1 if the station has a major park-ride lot, 0 if otherwise.
3) TR - an index equal to 1 if the station is a transfer point in the final Metro system, 0 if otherwise.

Station Number	Station	SS	PR	TR	Distance To Metro Center	% Owner	% Substandard Housing	% Non-White	Mean Income (1969 $)
1	Alabama Avenue	0	1	0	5.2	22	0	73	8,988
2	Anacostia	0	0	0	4.0	23	0	81	10,195
3	Archives	0	0	0	0.7		insufficient census data		
4	Benning Road	1	0	0	6.3	23	1	99	9,181
5	Brookland	1	0	0	3.7	45	1	60	11,652
6	Capitol Heights	0	0	0	7.7	25	1	100	9,125
7	Capitol South	0	0	0	2.0	26	4	52	15,057
8	Cleveland Park	0	0	0	3.3	28	1	42	26,480
9	Columbia Heights	0	0	0	2.7	18	4	93	7,944
10	Deanwood	1	1	0	5.8	36	3	100	8,324
11	DuPont Circle	0	0	0	1.3	6	4	25	12,914
12	Eastern Market	0	0	0	2.7	32	2	51	11,748
13	Farragut N/W	0	0	0	0.8	2	2	31	6,585
14	Federal Center SW	0	0	0	1.5	0	0	39	19,036
15	Federal City College	0	0	0	0.8	5	24	95	5,543
16	Federal Triangle	0	0	0	0.3		insufficient census data		
17	Foggy Bottom	0	0	0	1.4	11	2	6	20,932
18	Fort Totten	1	1	1	5.1	43	1	88	11,901
19	Gallery Place	0	0	1	0.3	5	13	34	8,169
20	Georgia Avenue	0	0	0	3.7	43	3	96	11,497
21	Judiciary Square	0	0	0	0.6	2	4	76	5,603
22	L'Enfant Plaza	0	0	1	1.2		insufficient census data		
23	McPherson Square	0	0	0	0.5	1	5	18	8,606
24	Metro Center	0	0	1	0.0	6	11	15	8,169
25	Minnesota Avenue	1	1	0	5.8	33	3	98	9,896
26	Navy Yard	0	0	0	2.2	3	3	94	5,021
27	Potomac Avenue	0	0	0	3.2	23	1	93	7,249
28	Rhode Island Avenue	1	1	0	2.8	32	1	95	9,987
29	Shaw	0	0	0	1.3	14	11	99	6,730
30	Smithsonian	0	0	0	0.7		insufficient census data		
31	Stadium/Armory	0	0	1	3.8	40	1	95	9,706
32	Takoma Park	1	0	0	6.9	63	1	94	12,828
33	Tenley Circle	0	0	0	4.9	71	1	3	19,710
34	U Street	0	0	0	1.5	18	8	99	7,460
35	Union Station	0	0	0	1.2	23	4	88	8,626
36	Van Ness - WTI	0	0	0	3.8	20	1	3	25,465
37	Waterfront	0	0	0	2.8	19	0	53	17,020
38	Zoological Park	0	0	0	2.5	81	2	4	19,580

TABLE 2.1 – Station Area Characteristics

In some cases, the census data was insufficient to obtain reasonable estimates of the housing stock and percent non-white population. These were severalstations in which there were very few residents and counts were either not published for disclosure reasons or were simply not meaningful. All parcels from such stations were omitted from the final sample.

Table 2.2 summarizes the same data as Table 2.1 but for all census tracts used in the study that were not part of the previously defined station areas. Only tracts from which parcels were sampled are included in the table. In each case, the transit station nearest to the census tract is listed for reference.

2.3 Auxiliary Data

Aside from the parcel and neighborhhod data, provided by the MPF, the Lusk directories and the Census, a number of other minor data sources were used as part of the data base. The following were employed in the final data set:

1) 1972 Regional Employment Census (conducted by Washington Council of Governements) -for retail and total employment on the cènsus tract level.

2) U.S. Census Maps -for measurement of Census tract areas and the distance from each sampled parcel to the nearest Metro station. This distance was:

Number	Tract/Station*	SS	PR	TR	Distance to Metro Center	% Owner	% Substandard Housing	% Non-White	Mean Income (1969 $)
39	7/Zoo Park	0	0	0	2.5	16	1	2	19,580
40	14/Tenley	0	0	0	4.9	51	1	5	18,825
41	80.01/Stadium	0	0	1	3.8	41	1	99	8,460
42	91.01/Rhode Island	1	1	0	2.8	71	1	85	12,611
43	18.02/Takoma	1	0	0	6.9	22	1	78	12,836
44	42.01/DuPont	0	0	0	1.3	6	6	84	7,609
45	43/U Street	0	0	0	1.5	10	4	83	8,926
46	73.04/Alabama	0	1	0	5.2	10	6	97	8,627
47	77.01/Anacostia	0	0	0	4.0	14	1	84	8,962
48	77.01/Benning Road	0	0	0	4.0	14	1	84	8,962
49	77.07/Benning Road	1	0	0	6.3	49	0	98	10,833
50	78.07/Capitol Hts.	1	0	0	7.7	26	1	100	8,595
51	79.01/Stadium	0	0	1	3.8	25	1	99	8,017
52	80.02/Stadium	0	0	1	3.8	31	2	97	8,587
53	81/Eastern	0	0	0	2.7	26	2	88	9,372
54	88.02/Union Station	0	0	0	1.2	37	1	98	8,864
55	89.02/Stadium	0	0	1	3.8	9	1	100	7,150

TABLE 2.2 - Characteristics of Tracts Outside Station Areas

* Tract indicates the census tract number
Station indicates the nearest Metro station

a) measured as the straight line distance from the middle of the street forming the boundary of the census block on which the parcel is located to the middle of the intersection where the Metro station is or will be located.

b) recorded in tenths of a mile, and interpreted as the parcel's being "no greater than" that distance from the nearest Metro station (i.e. - a parcel of 0.25 miles from the Metro station was recorded as 0.3, and was interpreted as no greater than 0.3 miles from the nearest station).

3) WMATA planning documents - for estimates of the opening dates for various stations in the WMATA system. Data on the number of years expected for the completion of the nearest Metro station was obtained from the original Metro construction schedule and its subsequent revisions. Table 2.3 shows the number of years to completion for various station areas at various times. The number of each years to completion for any transaction depended upon the station area and the date of the transaction. If a parcel transacted in a year of a construction schedule change, but before the change was announced, the parcel is assigned the number of years to completion based on the old schedule. If the transaction occurred during or after the month of the schedule change, the number of years to completion is that of the new schedule. For example, suppose a parcel in the Brookland station area was transacted in August, 1972. In this case, the number of years to completion assigned to it is 2 years, (3 years in 1971 minus 1 year for 1972). If the same parcel was purchased in December, 1972, the number of years until completion would be 4 years under the revised schedule.

4) Bureau of Labor Statistics - for housing and consumer price indices for Washington, D.C. These were used to deflate transaction prices to a constant dollar level. The base year used was 1969. The deflator was calcu-

Year of Transaction (+ Month of New Schedule)	Feb. 1969	1970	Jan. 1971	Nov. 1972	1973	Jul. 1974	1975	Aug. 1976
Station								
Alabama Ave	9	8	7	6	5	7	6	7
Anacostia	9	8	7	6	5	7	6	7
Archives	7	6	6	5	4	4	3	5
Benning Road	5	4	4	5	4	3	2	2
Brookland	4	3	3	4	3	3	2	1
Capitol Heights	7	6	6	6	5	6	5	1
Capitol South	5	4	4	4	3	2	1	1
Cleveland Park	5	4	5	5	4	4	3	4
Columbia Heights	7	6	6	6	5	7	6	7
Deanwood	5	4	4	5	4	3	2	2
DuPont Circle	3	2	2	2	1	1	1	1
Eastern Market	5	4	4	4	3	2	1	1
Farragut North	3	2	2	2	1	1	1	0
Farragut West	4	3	4	4	3	2	1	1
Federal Ctr, SW	5	4	4	4	3	2	1	1
Federal City Coll.	7	6	6	6	5	6	5	5
Federal Triangle	4	3	4	4	3	2	1	1
Foggy Bottom	4	3	4	4	3	2	1	1
Fort Totten	4	3	3	4	3	3	2	1
Gallery Place	5	4	4	4	3	2	1	1
Georgia Avenue	7	6	6	6	5	7	6	7
Judiciary Square	3	2	2	2	1	1	1	0
L'Enfant Plaza	5	4	4	4	3	2	1	1
McPherson Square	4	3	4	4	3	2	1	1
Metro Center	3	2	2	2	1	1	1	0
Minnesota Avenue	5	4	4	5	4	3	2	2
Navy Yard	9	8	6	5	4	4	3	7
Potomac Avenue	5	4	4	4	3	2	1	1
Rhode Island Ave	3	2	2	2	1	1	1	0
Shaw	7	6	6	6	5	6	5	5
Smithsonian	4	3	4	4	3	2	1	1
Stadium/Armory	5	4	4	4	3	2	1	1
Takoma Park	4	3	3	4	3	3	2	1
Tenley Circle	5	4	5	6	5	6	5	5
U Street	7	6	6	6	5	6	5	5
Union Station	3	2	2	2	1	1	1	0
Van Ness - WTI	5	4	5	5	4	4	3	4
Waterfront	7	6	6	5	4	4	3	7
Zoological Park	5	4	5	5	4	4	3	4

TABLE 2.3 - Number of Years to Completion

lated based on an average of the housing and overall prices for the Washington, D.C. metropolitan area. The deflator is as follows:

Year	Deflator
1966	0.88
1967	0.90
1968	0.94
1969	1.00
1970	1.06
1971	1.11
1972	1.17
1973	1.22
1974	1.36
1975	1.47
1976	1.60

Source: Handbook of Labor Statistics
U.S. Department of Labor, Bureau of Labor Statistics
Bulletin #1905
Washington, D.C.
U.S. Govt.Printing Office
029-001 - 01962-9

2.4 Summary of Samples

The three final data sets (single family, multi-family and retail) used in the model estimation were formed by integrating all of the above data into separate files, with each transaction augmented by station area and auxiliary information. These samples were then tabulated to check for coding and processing errors, and more significantly, to obtain some insight into the range of the data.

These tabulations (along with various hand-checking procedures) also uncovered a major flaw in the multi-family dwelling sample. Of the 771 multi-family apartments in the sample, 240 (approximately 31%) did not have a valid code for the number of dwelling units in the building. Rather than eliminate these observations, it was decided to estimate the number of dwelling units based on other properly coded variable values. This was done by regressing the number of dwelling units on some of the other variables, using only the 531 observations for which dwelling unit counts were available. The final model used for this purpose is as follows:

$$\text{number of dwelling units} = \underset{(1.50)}{3.15} + \underset{(.000103)}{.00204}\ (\text{lot area})$$

$$R^2 = .427$$

(The numbers below the coefficients are their respective standard errors.)

The equation was then used to predict the number of dwelling units for the remaining 240 parcels.

Tables 2.4, 2.5 and 2.6 are summaries of the major characteristics of the samples for owner-occupied dwellings, multi-family units and commercial establishments respectively. In each table, the sample mean, standard deviation, and range of the variables are given. In the case of the multi-family building sample, the statistics for the number of dwelling units include those parcels for which the value was estimated based on lot area.

Variable	Mean	Standard Deviation	Range
1. Deflated transaction price (1969 dollars)	$30,746	$17,954	$2,090-$124,150
2. Distance to nearest Metro station (miles)	0.3	0.6	0.1 - 1.2
3. Number of years until station completed	3.7	2.0	0 - 9
4. % dwellings owner occupied in station area	23.2%	15.69%	2 - 71%
5. % substandard dwellings in station area	3.3%	3.49%	0 - 11%
6. % non-whites in station area	59.6%	33.78%	2 - 99%
7. Annual income in station area	$12,255	$5,029	$5,021-$26,480
8. Distance to Metro Center (miles)	2.53	1.18	0 - 6.3
9. Parcel lot area (sq.feet)	1,917	1,297	532 - 15,168
10. Total employees per sq.mi. in station area	14,630	18,092	831 - 79,463
11. Retail employees per sq.mi. in station area	1,148	1,079	0 - 5,009
12. Population per sq.mi. in station area	27,220	12,439	6,217 - 49,675

Sample Size = 286

TABLE 2.4

Summary of Single-Family Dwelling Sample

Variable	Mean	Standard Deviation	Range
1. Deflated transaction price (1969 dollars)	$60,796	$192,404	$1,108 -$2.6 X 10^6
2. Distance to nearest Metro station (miles)	0.6	0.4	0.1 - 2.0
3. Number of years until station completed	3.7	2.1	0 - 9
4. % dwelling owner occupied in station area	25.1%	13.6%	2 - 81%
5. % substandard dwellings in station area	2.4%	3.3%	0 - 44%
6. % non-whites in station area	72.3%	31.1%	2 - 100%
7. Annual income in station area	$10,428	$3,410	$5,021-$26,480
8. Distance to Metro Center (miles)	0.4	0.3	0 - 7.7
9. Parcel lot area (sq.feet)	5,138	11,536	1,012 - 273,295
10. Total employees per sq.mi. in station area	9,020	19,155	617 - 266,983
11. Retail employees per sq.mi. in station area	1,006	1,332	0 - 18,646
12. Population per sq.mi. in station area	27,310	14,916	3,623 - 63,443

Sample Size = 771

TABLE 2.5

Summary of Multi-Family Building Sample

Variable	Mean	Standard Deviation	Range
1. Deflated transaction price (1969 dollars)	$54,422	$82,622	$500 - $829,472
2. Distance to nearest Metro station (miles)	.2	.2	0.0 - 1.4
3. Number of years until station completed	3.0	2.2	0 - 9
4. % dwelling owner occupied in station area	27%	13%	*
5. % substandard dwelling in station area	8%	27%	*
6. Annual income in station area	$9,322	$3,416	$5,021 - $26,480
7. Distance to Metro Center (miles)	1.92	1.40	0 - 6.3
8. Parcel lot area	2,879	6,090	613 - 56,016
9. Total employees per sq.mi. in station area	41,199	60,748	617 - 266,984
10. Retail employees per sq.mi. in station area	4,265	6,578	0 - 22,165
11. Population per sq.mi. in station area	24,625	16,021	2,207 - 63,442
12. Parcel floor area	2,520	2,862	400 - 19,589

Sample Size = 353

Note: 17% of the observations in this sample were transacted prior to 1969.

* not available

TABLE 2.6

Summary of Retail Establishment Sample

2.5 Evaluation of Samples

Prior to discussion the specific models which were developed, some general comments about the entire data base are appropriate. As is virtually unavoidable, compromises between total accuracy and practical considerations had to be made in the course of building the data base. Although the most damaging compromises were kept to a minimum within the time and budget constraints of the study, we were unable to correct all deficiencies of the data.

First, the parcel-specific data was taken from the 1974 Metropolitan Parcel File (MPF). The zoning classification, for example, could have been different in the year of a parcel's transaction. Since such information was not readily available (and also since we had little confidence in the zoning classification as an important explanatory variable in the models) we made no further attempt to correct this weakness.

Second, the sales, or transaction's prices, were not always entirely reliable. Inaccuracies or unreasonable figures were uncovered in the Lusk Directories. For example, prices at inconceivably low levels were reported next to reasonable prices for the same year. We suspected that many of the odd-looking prices were a reflection of the special practices of the real estate profession, such as transactions between household members.* Where a price appeared totally implausible, it was deleted from the samples.

Third, the general demographic data for the station areas was based entirely on the U.S. Census from 1970. As transactions occurred in six other years, there might be some cause to question the accuracy of 1970-based data. Despite some initial doubts, changes over the seven year period in question

*See Appendix A for a more complete discussion of potential reasons for these transactions.

were not deemed substantial enough to warrant revision of the demographic section of the data base.

Fourth, the sale's price contains information about people's valuation of any improvements which have been made in a parcel. Unfortunately, it was not possible to obtain data on improvements without searching through numerous building permits. Permit data was investigated, but found to be inaccessible in any machine-readable form.

Fifth, the area-specific data frequently misrepresented the actual environment in which a parcel existed. Neighborhood boundaries are extremely difficult to define, and parcels on the borders between station areas may well be assigned demographic characteristics which are not reasonable reflections of the actual conditions in a neighborhood.

Despite these deficiencies, the data base was deemed suitable (though not ideal) for estimating property value models. It is our belief that these deficiencies present no insurmountable problems; as in all empirical studies they simply make the final results less than perfect.

3. EMPIRICAL RESULTS

Each of the three data files described in the preceding chapter were used to estimate models of property transaction value. This chapter summarizes the results of these model estimation efforts. While the chapter is not a complete inventory of every single model estimated, an attempt is made to present the most important of the empirical results.

The following section is a brief discussion of the criteria used for model evaluation. In particular, a method developed by Box and Cox (1964) to test alternative functional forms is discussed. Section 3.2 defines the notation used in the later sections of the chapter, and provides some of the underlying rationale for why the less obvious of the variables were considered. Sections 3.3, 3.4 and 3.5 present the estimation results for models of single family houses, multi-family buildings and retail establishments respectively. Section 3.6 is a synthesis of the conclusions derived from the modelling efforts emphasizing the implications regarding the effect of transit in urban property values.

3.1 - Model Specification Process

There are at least two distinct approaches to the econometric modelling of urban property values. The first, a structural approach, would represent buyers' and sellers' behaviors in separate, but simultaneous, equations. The second approach implicitly or explicitly solves the simultaneous, structural equations fot a single equation in which price is a dependent variable and only variables which are not dependent variables in the structural equations are independent variables. This equation, part of the so-called reduced form, is often termed a hedonic price equation.

This study adopts the hedonic price approach for a number of reasons, the most relevant of which is that the development of a structural model of urban property values is a task which would require vastly more effort than the hedonic price approach. Moreover, the hedonic price approach can meet most of the objectives of the study.

The process of model specification on all three samples was largely an iterative one. Different variables and functional forms were tested, and either accepted or rejected. For each sample, however, at least one model was denoted as a "final", or best, specification. (In some cases more than one model was selected for later use.) The choice of which specification (or specifications) was "best" was based on:

(1) statistical considerations - goodness-of-fit and statistical significance of the estimated parameters;

(2) forecasting tests - alternative models were used to forecast property values for all samples in selected station areas, and forecasts were compared with observed transactional values both at the low and high ends of the values' range.

(3) analysis of functional form - a statistical procedure developed by Box and Cox was used to test the appropriateness of alternative functional forms.

This last technique requires some elaboration. Suppose one is considering a particular variable y in a regression model $f(y) = X\beta + \varepsilon$. However, suppose that one is uncertain as to whether the specification of f(y) should be linear, logarithmic, quadratic or some other form*. Box and Cox examine a "family" of transformations with the following specification:

---

* The Box and Cox method is actually much more general and can be applied to both dependent and independent variables. The description given here is limited to how the technique was used in this study.

$$y^* = \begin{cases} \dfrac{y^{\lambda}-1}{\lambda\, gm(y)^{\lambda-1}} & \text{if } \lambda \neq 0 \\[2ex] \dfrac{\log y}{gm(y)} & \lambda = 0 \end{cases}$$

where: $gm(y)$ is the geometric mean of y

λ is a parameter to be estimated

The parameter λ allows for representation of a wide range of functional forms. (Note that as $\lambda \rightarrow 0$ the expression $y^{\lambda-1}/\lambda gm(y)^{\lambda-1}$ converges to $\log\lambda/gm(y)$.) If $\lambda = 1$, then the form of the regression is linear; if $\lambda = 0$, then it is log-linear. Moreover, "intermediate" function forms such as $y = (X\beta)^{\frac{1}{2}}$ can be represented or by value of λ equal to 2.

The value of λ can be estimated either directly by maximum likelihood or by searching over different values of λ, estimating the equation $f(y) = X\beta$ by ordinary least squares for each value and choosing the λ which minimizes the sum of the squared residuals. The latter technique has the distict advantage of not requiring a non-linear model estimation procedure, and was therefore used in this study.*

---

* The two techniques are equivalent if the disturbances are independent, homoskedastic and normally distributed.

3.2 - Definition of Variables

The response or dependent variable of interest is, of course, the transaction price of a parcel, deflated to 1969 dollars. In order to understand the variation of this price from parcel to parcel, we defined a set of explanatory variables to be included in the regression models. These explanatory variables can be usefully grouped into three categories: transit system-related, demographic, and parcel-specific. While the names and a short description of each variable are given in Table 3.1, a brief commentary on some of these is appropriate. In effect, the inclusion of each variable represents an hypothesis about the way in which property buyers and sellers evaluate market conditions in Washington, D.C.

The first group of variables has, naturally, the most direct bearing on transportation policy since changes in them could affect the distribution of land values and ultimately land uses. Proximity to a Metro station would be expected to increase a parcel's price as most people would enjoy the convenience of not having to walk great distances or first take a bus before riding the subway. In addition to "DIST" we also defined a dummy variable (PXDUM) to capture any negative effects of being too close to a station. That is, we hypothesized that the value of those parcels within 0.1 (or 0.2) miles would be degraded to some extent because of increased pedestrian and vehicular traffic and the resulting noise and air pollution increases. Similarly, we defined dummy variables for a station's being above ground and for its having park-ride facilities, so that any negative impacts associated with these characteristics could be included in the regression. Because the Metro system had not yet been completed it seemed obvious that those parcels whose nearest station was about to be opened would tend to be more influenced than those whose station's opening was much further in the future. This is reflected in the variable YR , the number of years until the station is opened.

NAME	VARIABLE TYPE AND DEFINITION	
	DEPENDENT VARIABLE:	
DPRICE	TRANSACTION PRICE, DEFLATED TO 1969 DOLLARS	
	EXPLANATORY VARIABLES	
GROUP ONE: TRANSIT SYSTEM RELATED		
DIST	Straight line distance to nearest Metro station (in miles)	
PXDUM	Proximity dummy:	1 if parcel is located within a specified distance to station 0 otherwise
YR	Number of years to completion	
SS	Dummy variable:	1 if station is above ground 0 otherwise
PR	Dummy variable:	1 if station has park-ride lot 0 otherwise
GROUP TWO: DEMOGRAPHIC		
OWNER	% owner-occupied dwellings	
HOUSE	% substandard housing	
NWHITE	% non-white	
INCOME	Mean income ($/yr)	
EDEN	Total employment density (employees/sq.mi)	
RDEN	Retail employment density (employees/sq.mi)	
PDEN	Population density (persons/sq.mi)	
GROUP THREE: PARCEL SPECIFIC		
DISTM	Distance by transit to Metro Center (in miles)	
LOTA	Lot area (sq.ft.)	
FLAREA	Floor area of parcel's improvement (sq.ft.)	
TOTV	Total assessed valuation ($)	
LANV	Assessed valuation of land ($)	
ZDUM	Zoning dummy:	1 if parcel's zoning class and property type are identical 0 otherwise
CBDDUM	CBD dummy:	1 if parcel is located in CBD 0 otherwise
NBDU	Number of dwelling units	
DUDUM	Dwelling unit dummy:	1 if NBDU was actually recorded for observation 0 otherwise
PRE	Pre-1969 dummy:	1 if transaction was pre'69 0 otherwise

TABLE 3.1 Variables in the Models

Note that there is no measure of highway access included in Table 3.1. This is because the study area was limited to the city of Washington, where variations in average travel speed are far smaller than variations in transit access. If the study had been extended to include the sections of Maryland and Virginia which comprise the metropolitan area, some auto level of service measure would probably have been essential.

In the second group we defined variables to represent demographic characteristics of the neighborhood in which a parcel was located. It is well known in the real estate market that the environment surrounding a property often plays a major role in determining the final price regardless of the condition of the parcel in question. While the seven variables we defined are by no means exhaustive, they represent most of the major dimensions along which a property's environment could reasonably be measured. Mean income and the percent substandard housing should be straightforward. The percent of owner-occupied dwellings in an area, should probably have a depressing effect on the values of single family dwellings. Its effect on multiple family structures is somewhat less clear, since renters may or may not seek areas with predominantly rental housing. The percent non-white was, from the beginning, a variable about which no a priori expectation was formed; nevertheless, it could play some part in influencing the final market price. Likewise, it was not clear how the density-related variable would affect prices in specific markets.

The third group of variables corresponded to hypotheses about the attributes of specific parcles independent of who lived there or their proximity to the Metro system. The rationale behind any of these (e.g. lot area, structure's condition) should be obvious and need not be explained. The dummy variables, however, are not as straightforward. In specifying the zoning variable we hypothesized that consistency of zoning class and property type was viewed in the marketplace as desirable. Parcels located in the central business district (CBD) enjoy proxmity to at least two of Metro's stations (i.e. within ½ mile), and should therefore have greater value than equivalent parcels outside. In using data on multi-family dwellings, we were forced to supply an estimated number of dwelling units for those parcels whose records were missing the appropriate data. Consequently, we defined a dummy variable to account for this fact (DUDUM). In the case of the models for retail store prices, information from 1966 to 1968 was also used; consequently, a dummy variable PRE measured any differences between the '66-'69 and '69-'76 intervals.

3.3 - Single Family Dwelling Models

a) Linear Additive Model

The primary regression run in the linear additive form with deflated price as the dependent variable produced the following results:

(3.3.1) DPRICE = 37776 - 2480 DIST + 4604PXDUM - 951 YR - 5938 SS
(1.90) (1.50) (2.11) (1.03)

+ 18,870 PR - .7630 DISTM - 13,299 CBDDUM
(2.37) (3.55) (1.02)

+ 97.4 OWNER - 2224 HOUSE + 0.253 NWHITE + 0.725 INCOME
([illegible].28) (4.38) (0.01) (1.78)

+ 0.025 EDEN + 2.835 RDEN + 0.0086 PDEN
(0.38) (2.47) (0.39)

+ 4.210 LOTA + 6733 ZDUM $R^2 = 0.598$
(6.13) (2.13)

(t-statistics given within the parentheses)

These results suggested the NWHITE, EDEN and PDEN variables (corresponding to the percent non-white, employment density and population density respectively) be examined more closely. Four subsequent regression runs were done, three dropping out each of the variables individually, and one dropping out all three together. When one of the variables was dropped out individually, the other variables generally did not change magnitude or significance in a dramatic manner.

Looking at these variables individually, it was observed that total employment density and retail employment density were highly correlated, with a partial correlation coefficient of 0.657. Due to the higher significance of RDEN in the model, and the correlation between it and EDEN (employment density), a decision was made to include RDEN in the model while dropping EDEN.*

Population density, surprisingly, was not significant in any of the three runs. It was expected that transaction prices would be negatively associated with population density, i.e. the higher the population density, the lower the transaction price would be. However, the coefficient was positive in all three cases, with a low magnitude, and a t-statistic of approximately 0.4. The PDEN variable contributes a small portion of the transaction price as estimated by the linear model, at most a contribution of $1,000. Due to these results, a decision was made to drop the PDEN variable from consideration in the model.

Percent non-white (NWHITE) was another insignificant variable in these runs. This is not surprising since the evidence regarding whether or not non-whites pay more or less for equivalent housing is unclear. Due to the low significance, low magnitude of this variable and the lack of a convincing conceptual model to describe the relationship between the composition of a neighborhood and the value of a property, a decision was made to drop the NWHITE variable from the model.

---

* This results in a biased estimate for the coefficient of RDEN, and some care should be taken to avoid policy conclusions based on changes in RDEN when using the models.

The regression model without the NWHITE, EDEN, PDEN variables yielded the following results:

$$(3.3.2) \quad \text{DPRICE} = 37{,}758 - \underset{(2.10)}{10{,}235\ \text{DIST}} + \underset{(1.04)}{3279\ \text{PXDUM}} - \underset{(2.52)}{1098\ \text{YR}} - \underset{(1.39)}{8074\ \text{SS}}$$

$$+ \underset{(2.60)}{20{,}215\ \text{PR}} - \underset{(3.23)}{6724\ \text{DISTM}} - \underset{(0.86)}{10{,}600\ \text{CBDDUM}}$$

$$+ \underset{(1.38)}{90.5\ \text{OWNER}} - \underset{(4.28)}{2103\ \text{HOUSE}} + \underset{(2.79)}{0.757\ \text{INCOME}}$$

$$+ \underset{(3.66)}{3.136\ \text{RDEN}} + \underset{(6.33)}{4.226\ \text{LOTA}} + \underset{(2.34)}{7205\ \text{ZDUM}}$$

$$R^2 = 0.600$$

(The change in magnitude of coefficient of DIST (distance to station) was partially due to the correction of a transcription error in the distance variable.) With this new model (Eq. 3.3.2) as a base, the next series of regression runs focused on the Metro related variables.

A reciprocal form for the distance to station variable was tried. This functional form reflects the hypothesis that the linear form understates the effect of proximity to a station for nearby and distant parcels overstates it for the middle range. The PXDUM variable was introduced to capture any special effects of a parcel being extremely close to the station. Thus, the measurement of the distance to nearest Metro station influence is:* αDIST1 + β*PXDUM where DIST1 = 1/(DIST+1).

*The shifting of distance in the denominator was to avoid the possibility of dividing by zero.

The PR variable was modified to measure the effect of only those parcels within walking distance (arbitrarily chosen distance of 0.4 miles) of a park-ride station.

Various models were estimated using the reciprocal for distance, with and without the PR, CBDDUM, and PXDUM variables. The results of these runs confirm the intuitive judgement that the reciprocal form for distance explains the decline of housing price with distance to station better than the linear form (t-statistic of 3.1 as compared to 2.1). The PXDUM coefficient was small and statistically insignificant with the reciprocal form for distance, which suggests the hypothesis that the reciprocal form seemed to "pick up" a substantial amount of the close proximity effect, while the linear form represented a mis-specification for short distances.

The modified park-ride variable proved to be of little use as none of the observations in the sample were within walking distance (0.4 miles) to a park-ride station. Since the purpose of this variable was to measure any nuisance effects of a park-ride station, inclusion of the PR variable would have produced misleading results. Consequently, the variable was dropped from the model.

The elimination of the PR dummy variable resulted in the decrease of the statistical significance of the SS (above surface dummy) variable. Closer examination of the two variables uncovered a partial correlation of 0.55, due principally to the inclusion of only two METRO stations in the sample which were at-grade, one of which contained a park-ride facility.

The CBDDUM (central business district dummy) variable was consistently insignificant in the models (t-statistic of 0.6). In addition, the area defined as the CBD for the CBDDUM variable contained a small number of single family dwellings. As a result, this variable was dropped from the model.

The final form of the linear additive model, with both the SS and PR variables dropped is:

$$
\begin{array}{lllll}
(3.3.3) & DPRICE = 24{,}988 & + 1212\ DIST1 & - 876\ YR & - 6192\ DISTM \\
 & & (3.06) & (2.09) & (5.03) \\
 & & + 73.7\ OWNER & - 2023\ HOUSE & + 0.899\ INCOME \\
 & & (1.22) & (5.45) & (3.90) \\
 & & + 3.533\ RDEN & + 3.986\ LOTA & + 4492\ ZDUM \\
 & & (4.29) & (6.19) & (1.68)
\end{array}
$$

$$R^2 = 0.592$$

Another final form of the model was run with the SS variable included but transformed to represent the effect of only those parcels within 0.5 miles of a Metro station which is a grade. In this model form, the SS variable measures the nuisance effect of an above ground station on parcels within close proximity to it. The model run produced a coefficient for the SS variable of -10,500 with a t-statistic of 2.07. A decrease of $10,500 in property value when a parcel is within 0.5 miles of an above ground station seemed unreasonably high. A closer examination of the sample revealed that only 10 observations were within 0.5 miles of an above ground station, one of the stations not to be completed for at least 4 years.

Consequently, additional observations were selected to be added to the sample. Transacted parcels within 0.4 miles of the Rhode Island Avenue, Brookland, and Fort Totten stations were obtained. These particular stations were chosen because they were already opened (Rhode Island Avenue) or were expected to be opened within a year (Brookland, Fort Totten), and they were all above grade stations, two of which (Rhode Island Avenue, Fort Totten) had park-ride facilities. A total of 40 observations were obtained and added to the previous sample.

b) Mixed/Linear Log Model

The preliminary regression run using a product form model (or log-log model) with the log of deflated price being the dependent variable produced the following results *:

$$\begin{aligned}
(3.3.4)\quad \ln(\text{DPRICE}) = {} & -1.5021 - \underset{(1.03)}{0.316\ln(\text{DIST})} - \underset{(4.42)}{0.246\ln(1+\text{YR})} \\
& - \underset{(2.97)}{0.60\ \text{SS}} + \underset{(1.74)}{0.513\ \text{PR}} + \underset{(0.97)}{0.091\ln(1+\text{OWNER})} - \underset{(1.17)}{0.117\ln(1+\text{HOUSE})} \\
& + \underset{(1.30)}{0.085\ln(1+\text{NWHITE})} + \underset{(4.24)}{0.839\ln(\text{INCOME})} + \underset{(0.02)}{0.0071\ln(\text{DISTM})} \\
& + \underset{(3.24)}{0.224\ln(\text{LOTA})} + \underset{(4.99)}{0.222\ln(\text{EDEN})} - \underset{(1.67)}{0.0303\ln(\text{RDEN})} \\
& + \underset{(0.39)}{0.0299\ln(\text{PDEN})} + \underset{(0.30)}{0.0346\ \text{ZDUM}} + \underset{(1.98)}{0.248\ \text{PXDUM}} \\
& + 0.250\ \text{CBDDUM}.
\end{aligned}$$

$$R^2 = 0.599$$

Given the information supplied by both the linear additive forms (Eqs.3.3.1, 3.3.2, 3.3.3) and the log-log form, a mixed linear/log form was specified using the functional form which implies the most plausible causal relationship. In some cases where no intuitive basis for choice of functional from existed goodness of fit (as measured by low standard errors in earlier equations) was used as a criterion to confirm specific hypotheses. The table on the following page summarizes the functional form chosen for each variable and a brief explanation of why that particular form was chosen.

---

*All of the independent variables which include possible zero values were translated by adding one in the logarithmic forms. All variables denoted by ln are natural logarithms.

Table 3.2 - Choice of Variables for Single Family Dwelling Model

VARIABLE	FUNCTIONAL FORM CHOSEN	EXPLANATION
DIST	Reciprocal form	-Implies a behavioral hypothesis that parcels closest to the station benefit the most, and that increasing distance rapidly decreases the benefits.
YR	Log form	-Implies behavioral hypothesis that people will value property more when the anticipated completion date is close at hand, and that this effect will level off when years to completion are large.
SS	Dummy variable	-This variable is defined exactly as in the linear additive model.
OWNER	Log form	-The log form was chosen under the hypothesis that there are diminishing marginal benefits to the percentage of owner-occupied dwellings in the neighborhood.
HOUSE	Log form	-The log form was chosen under the behavioral hypothesis that the first few substandard units would cause the greatest decline in property values. As the neighborhood becomes more blighted, the decline in value becomes smaller as additional substandard units are less
INCOME	Log form	-The hypothesis is that the mean income of the neighborhood is more important in determining property values when it is at an intermediate low-middle income range than at a high income range. -The log form provides a better fit than the linear form (t-statistic of 6.0 as compared with 3.5)
DISTM	Linear form	-The linear form provides the best fit.
LOTA	Linear form	-A linear form was chosen because the amount of land area should be proportional to the value of the land. -Both linear and log form fit the data well.
RDEN	Linear form	-The linear form consistently provides the best fit.
ZDUM	Dummy variable	-This variable was generally insignificant in the mixed form. However, it was included in the model due to its significance in the linear form.

The final form of the mixed linear/log model with both the SS and PR variables dropped is:

(3.3.5)
$$\text{DPRICE} = -117{,}060 + \underset{(2.52)}{1001}\ \text{DIST1} - \underset{(2.93)}{4210}\ \ln(1+\text{YR}) - \underset{(5.07)}{6104}\ \text{DISTM}$$
$$+ \underset{(2.47)}{4573}\ \ln(1+\text{OWNER}) - \underset{(4.48)}{7512}\ \ln(1+\text{HOUSE}) + \underset{(5.74)}{15{,}624}\ \ln(\text{INCOME})$$
$$+ \underset{(4.96)}{4.57}\ \text{REDN} + \underset{(6.52)}{4.139}\ \text{LOTA} + \underset{(.87)}{2462}\ \text{ZDUM}$$
$$R^2 = 0.601$$

c) Box and Cox Models

A Box and Cox analysis was performed to determine the best functional form for the model. These models were estimated with the transformed price using various values for λ (see section 3.1). The specification used was the final linear form without the SS variable (Eq.3.3.1). Values for λ included -1, -0.5, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0 and 2.

The results of these regressions indicate that a value for λ of 0.5 produced the regression with the best fit, i.e. the smallest sum of the squared residuals. The value of 0.5 for λ indicates that the best functional form for the model is one between a pure linear ($\lambda = 1$) and pure log forms ($\lambda = 0$). Statistical tests indicated that the estimated value of $\lambda = .5$ differed from both 0 and 1 at the 95% confidence level. Due to these results, a third functional form of the model, the Box and Cox form, was adopted for later use. The final Box and Cox model is summarized below:

(3.3.6) $$DPRICE = (0.003147 \cdot REGR\_EXP + 1)^2$$

where REGR_EXP =

$$= \underset{(8.79)}{49{,}862} + \underset{(3.00)}{1077} \cdot DIST1 - \underset{(2.14)}{814} \cdot YR - \underset{(5.18)}{5767} \cdot DISTM$$

$$+ \underset{(1.72)}{93.8}\ OWNER - \underset{(6.20)}{2{,}084}\ HOUSE + \underset{(4.09)}{0.852} \cdot INCOME + \underset{(4.44)}{3.313} \cdot RDEN$$

$$+ \underset{(4.37)}{2.58} \cdot LOTA + \underset{(1.94)}{5054} \cdot ZDUM$$

$$R^2 = 0.603$$

3.4 - Multi-family Building Models

a) Summary of Key Models

Although many models of multi-family building values were estimated, a discussion of four of them suffices to give the reader insight into our results. Although we experimented with linear model formulations as well as inverse specifications for the "distance to station" variable, we always found the predominantly logarithmic formulation to produce the most reasonable results*. The first model was estimated with all variables which were considered. The resulting equation was as follows:

(3.4.1)
$$\ln(DPRICE) = \underset{(.175)}{0.283} - \underset{(-2.34)}{0.48}\ln(DIST) - \underset{(-2.12)}{0.110}\ln(YR+1)$$
$$- \underset{(-1.49)}{0.119}(PR) + \underset{(2.56)}{0.191}\ln(DISTM)$$
$$- \underset{(-3.64)}{0.189}\ln(OWNER+1) - \underset{(-3.90)}{0.0047}(NWHITE)$$
$$+ \underset{(2.82)}{0.072}\ln(HOUSE+1) + \underset{(4.31)}{0.686}\ln(INCOME)$$
$$- \underset{(1.42)}{0.067}\ln(EDEN) + \underset{(3.01)}{0.133}\ln(PDEN) \qquad R^2 = .618$$

*At least in the case of the dependent variable, this judgement was confirmed by the Box and Cox models described below.

(3.4.1) (cont.)

$$\begin{aligned}
&- \underset{(-1.81)}{0.038} \ln(\text{RDEN}) + \underset{(3.01)}{1.083} (\text{CBDDUM}) \\
&+ \underset{(0.15)}{0.013} (\text{ZDUM}) + \underset{(0.53)}{0.045} (\text{PXDUM}) \\
&+ \underset{(2.68)}{0.161} \ln(\text{LOTA}) + \underset{(9.02)}{0.605} (\text{DUDUM}) \\
&+ \underset{(10.96)}{0.491} \ln(\text{NBDU})
\end{aligned}$$

The signs on all of the coefficients (with the possible exception of HOUSE, the percent of substandard dwellings) of this model are reasonable, but indicate very low levels of statistical significance for several variables, especially the qualitative, summy variables. The theoretical doubtfulness of these dummy variables led to a more restricted form as follows:

(3.4.2)

$$\begin{aligned}
\ln(\text{DPRICE}) = &- \underset{(-3.15)}{0.492} - \underset{(-3.88)}{0.187} \ln(\text{DIST}) - \underset{(-2.15)}{0.111} \ln(\text{YR}+1) + \underset{(0.011)}{0.137} \ln(\text{DISTM}) \\
&- \underset{(-4.15)}{0.206} \ln(\text{OWNER}+1) - \underset{(-3.68)}{0.0043} (\text{NWHITE}) + \underset{(3.68)}{0.084} \ln(\text{HOUSE}+1) + \underset{(6.01)}{0.829} (\text{INCOME}) \\
&- \underset{(-1.24)}{0.022} \ln(\text{RDEN}) + \underset{(2.54)}{0.144} \ln(\text{PDEN}) + \underset{(3.60)}{1.233} (\text{CBDDUM}) + \underset{(2.43)}{0.142} \ln(\text{LOTA}) \\
&+ \underset{(9.20)}{0.615} (\text{DUDUM}) + \underset{(11.24)}{0.500} \ln(\text{NBDU})
\end{aligned}$$

$$R^2 = .616$$

In this functional form, the PR (park ride dummy), EDEN (employment density), ZDUM (a zoning classification dummy) and PXDUM (a proximity to station dummy) were omitted. Employment density was omitted largely because it represents the same causal factor as retail density. This change produced a somewhat unexpected coefficient of retail density that was lower than either of the previous coefficients of employment or retail density. While the t-statistic for the coefficent in the second model is lower than either of the two previous ones, the reliability of the coefficient of RDEN (as indicated by its small standard error) is actually greater. For this reason, the EDEN was omitted from later models. To test the validity of using the parcels' deflated transaction prices as the dependent variable, we also ran one of the models using the logarithm of assessed total value of a parcel denoted as TOTV. The model (with a somewhat reduced set of variables as compared with equation 3.4.2) is as follows*:

$$(3.4.3)\quad \ln(\text{TOTV}) = \underset{(9.15)}{4.72} - \underset{(-.51)}{.02}\ln(\text{DIST}) - \underset{(-.21)}{.008}(\text{YR}+1) - \underset{(-2.33)}{.30}(\text{PR}) + \underset{(1.57)}{.05}(\text{DISTM})$$

$$- \underset{(-3.38)}{.14}\ln(\text{OWNER}+1) - \underset{(-7.65)}{.001}(\text{NWHITE}) - \underset{(-.83)}{.001}(\text{HOUSE}+1) + \underset{(3.61)}{.08}\ln(\text{EDEN})$$

$$- \underset{(-2.50)}{.09}\ln(\text{PDEN}) + \underset{(25.01)}{.96}(\text{DUDUM}) + \underset{(39.89)}{.84}\ln(\text{NBDU})$$

$$R^2 = .794$$

The results seemed to underscore the factors which assessors probably include in their evaluations of property value. One of the most striking aspects of the model based on assessed value rather than transacted value is that the coefficient of the logarithm of the distance to the nearest station in the two models differ by a factor of more than 9. In equation 3.4.2, the estimated coefficient for the log of distance is -.19 (with a standard error of .049), while in the equation using assessed value, the corresponding coefficient is only -.02 (with a standard error of .04). Similarly, the coefficient of the number of years before the station is open is far less in the assessed value model than in the transaction value model (-.11 as compared to -.008).

*This model was estimated at an earlier stage in the model development process, and time and budget did not allow for re-estimation using the specification in equation 3.4.2. All of the models using assessed value produced similar results, and the conclusions below would not be likely to change if the model were re-estimated.

Apparently, assessors tend to significantly understate the effect of transit systems on multi-family residential structures. They appear to use criteria which are more closely linked to "easy to observe" factors such as the number of dwelling units (NBDU). Furthermore, the goodness-of-fit for the assessed value model (as measured by the value of R^2) is much greater than that obtained in any of the models based on transaction prices. This would indicate that market prices are subject to a wide range of influences that assessors may understate or totally ignore.

b) Box and Cox Models

As with the single family dwelling model, a series of Box and Cox analyses of transformations was performed. For the model in equation (3.4.2), the estimate of λ in the transformed dependent variable was .05, indicating a functional form very close to logarithmic. Statistically, this value of λ was not different from 0 at reasonably confidence levels. It was therefore concluded that the logarithmic form in equation (3.4.2) was a suitable transformation.

The final Box and Cox estimates are given below*.

$$\frac{(\text{DPRICE})^{.05} - 1}{.05\,(\text{gm}(\text{DPRICE}))^{-.95}} =$$

$$(3.4.4)\quad \underset{(-1.79)}{-\,98364} - \underset{(-3.82)}{6464} \ln(\text{DIST}) - \underset{(-2.16)}{3914 \ln(\text{YR+1})} + \underset{(1.89)}{4556 \ln(\text{DISTM})}$$

$$\underset{(-4.15)}{-\,7271 \ln(\text{OWNER+1})} - \underset{(-3.65)}{149\,(\text{NWHITE})} + \underset{(3.57)}{2877 \ln(\text{HOUSE+1})} + \underset{(6.03)}{29265 \ln(\text{INCOME})}$$

$$\underset{(-1.08)}{-\,658 \ln(\text{RDEN})} + \underset{(2.46)}{4910 \ln(\text{PDEN})} + \underset{(3.73)}{44953\,(\text{CBDDUM})} + \underset{(2.36)}{4859 \ln(\text{LOTA})}$$

$$+\,\underset{(9.26)}{21775\,(\text{DUDUM})} + \underset{(11.79)}{18445 \ln(\text{NBDU})}$$

$$R^2 = .62$$

*This model can be put in the same general form as equation 3.3.8 by noting that the geometric mean of the deflated housing price in the multi-family building sample is \$ 35,138.

3.5 - Retail Establishment Models

a) Preliminary Models

The process of model development for retail establishment was nearly identical to that for single and multi-family buildings. An initial set of exploratory specifications were used to limit the later scope of the development process. These initial runs included all of the independent variables, and suggested the following:

1) that given the available data, it would be impossible to obtain a reliable estimate of the coefficient of PR (park-ride station dummy variable). This was because only 10% of the observations were near park-ride stations.

2) the coefficients for HOUSE (percent substandard dwellings) and NWHITE (percent non-white) consistently had both very small magnitudes and high standard errors.

3) that the empirical results from the log-log models were generally superior to alternative functional forms.

The second stage models developed a log-log from model in which most of the variables had coefficients which were significantly different from zero at the 95% level of confidence. This model is as follows:

$$\begin{aligned}(3.5.1)\quad \ln(\text{DPRICE}) = {} & \underset{(5.70)}{6.51} - \underset{(-1.50)}{.54}\ln(\text{DIST}) - \underset{(5.26)}{.55}\ln(\text{YR}+1) - \underset{(-2.22)}{.45}\,\text{SS} \\ & + \underset{(1.95)}{.38}\ln(1+\text{RENTER})^{*} - \underset{(-1.24)}{.07}\ln(\text{INCOME}) + \underset{(.65)}{.02}\ln(\text{DISTM}) \\ & + \underset{(6.07)}{.56}\ln(\text{LOTA}) + \underset{(3.43)}{.26}\ln(\text{FLAREA}) + \underset{(1.71)}{.05}\ln(\text{RDEN}) \\ & - \underset{(-3.67)}{.80}\text{PRE} - \underset{(-4.68)}{.30}\ln(\text{PDEN}) \\ & \qquad R^2 = .506\end{aligned}$$

* In these models, the percent of nearby households renting their dwelling (RENTER) was used instead of the percent owner-occupied. Note that RENTER + OWNER ≠ 1, since some units were vacant when the census was taken.

All the coefficients besides those for INCOME and DISTM had the expected signs* .

b) <u>Model Refinements</u>

In the next stage of model development, RENTER, SS, RDEN, EDEN, INCOME, DISTM and CBDDUM were included singly and in combination with each other. In addition, four new variables, ratios of the employment and density data for properties both inside and outside the CBD, were tested. Finally, the log of the difference between lot area and floor area, was introduced as a proxy for available parking lot space. Then, in cases where floor area was greater than lot area, a dummy variable, HIGHRISE was included as a proxy for the existence of multiple stories in the structure.

The ratio of retail employment density to population density (RDEN/PDEN) was intended to be a measure of the supply of retail stores per person. This variable was segmented into two variables corresponding to CBD and non-CBD parcels. (These are denoted below as R/PCBD and R/PNCBD for within and outside the CBD respectively). Two alternative effects were possible: (1) a competition effect, where the higher the ratio (*i.e.*, the higher the number of retail employees per person , the lower the transaction price of the retail property) and (2) an agglomeration effect, where the higher the number of retail stores in the area, the more business would be generated and the higher the transaction price of the retail property.

---

* It was later discovered that INCOME and DISTM were incorrectly coded in this sample. This was corrected in all subsequent models.

As the CBDDUM, EDEN, and DISTM (CBD dummy, employment density and distance to Metro Center) variables were used in combination and omitted, the coefficients of RDEN/PDEN in the CBD were consistently positive and ranged from .24 to .73 with very high t-statistics. These results were in accord with the agglomeration effect which was expected to take place in the CBD. The corresponding non-CBD coefficients were all positive, but smaller in value than their corresponding CBD counterparts by about a factor of ten. The inclusion of EDEN and DISTM had very large effects on the values and significance of the non-CBD effect; the values dropped dramatically and as a result, the coefficients became almost totally insignificant*.

Thus, agglomeration effects were found to be greater than competitive supply effects. More importantly, the separation of CBD and non-CBD samples resulted in very different estimates for the variables.

The ratio of non-retail employment density to retail employment density, (EDEN-RDEN)/RDEN, is a measure of the number of employees not involved in retail sales who generate retail business during working days, particularly during lunch hours 11:00 am to 2:30 pm; it provides some indication of the level of the potential shoppers who might find a retail establishment conveniently located nearby. Furthermore, it is also a measure of the competitive supply of retail establishments in the area. As with the retail employee to population ratio, EDEN-RDEN/RDEN, this variable was divided in order to allow for different coefficients within and outside the CBD. The resulting variables were denoted as QCBD and QNCBD respectively. Again, there were two alternative effects which may have been explained by the regression:

(1) The variables might have large positive coefficients, signifying

---

* When both EDEN and DISTM were not included, the coefficient value for R/PNCBD was .17, the t-statistic was 5.26. When both were included, the coefficient value was .002, and the t-statistic fell to .03.

that the higher the number of potential shoppers, the higher the transaction price of the retail property, and (2) that the variables might have negative coefficients, signifying that there are agglomeration benefits of locating among a large number of other retail stores and/or negative effects of locating among a large number of non-retail establishments.

Within the CBD, the coefficient of the ration (EDEN-RDEN) /PDEN variables was consistently positive and ranged in value from .21 to 1.00, with t-statistics which ranged from (2.16) to (4.63). The non-CBD variables, like the non-CBD ratio PDEN/RDEN variables, had coefficients which were all considerably lower in absolute value, and in many cases, negative. Again, EDEN and DISTM (employment density and distance to Metro Center) had insignificant coefficients. The negative signs were suggestive of the hypothesis of significant agglomeration effects, but the high standard errors cast some doubt as to their reliability as proof of that hypothesis*.

A third type of variable, the difference between lot area and floor area was used as a proxy for parking space availability. (This variable was denoted as PKLOT.) The interpretation of the variable could also include pedestrian plazas or simply wider sidewalks which might give the retail establishment a more spacious and appealing appearance to customers. In many samples, lot area was found to be less than floor area, indicating that the retail store was more than one story. In such cases, PKLOT was assigned a zero and the effect of a multi-stories structure on the dependent variable was picked up by a dummy variable, HIGHRISE, defined as 1 when PKLOT was zero.

The variable PKLOT was used in conjunction with FLAREA (floor area); LOTA (lot area) was dropped since it basically explained the same effect.

---

* Chow tests were performed to test whether the split into CBD and non-CBD effects yielded significantly different results than pooling the coefficients. It was possible to reject the null hypothesis of pooled coefficients at the 95% confidence levels for both PDEN/RDEN and (EDEN-RDEN)/PDEN.

The final model estimated using these new variables is as follows:

$$\begin{aligned}(3.5.2)\quad \ln(\text{DPRICE}) = & -.544 - .678 \ln(\text{DIST}) - .200 \ln(\text{YR}+1) \\ & \;\;(-.38)\quad\;\; (-2.01) \qquad\qquad (-1.87) \\ & + .675 \ln(\text{INCOME}) + .153 \ln(1+\text{PKLOT}) + .562 \ln(\text{FLAREA}) \\ & \qquad (4.52) \qquad\qquad\qquad (4.19) \qquad\qquad\qquad (8.35) \\ & - .328\ \text{PRE} + .399 \ln(\text{R/PCBD}+1) + .107 \ln(\text{R/PNCBD}+1) \\ & \quad (-1.57) \qquad (5.33) \qquad\qquad\qquad\qquad (2.78) \\ & + .342 \ln(\text{QCBD}+1) + .092 \ln(\text{QNCBD}+1) + .533\ \text{HIGHRISE} \\ & \qquad (4.75) \qquad\qquad\qquad (1.78) \qquad\qquad\qquad (2.07)\end{aligned}$$

$$R^2 = .5588$$

All the coefficients had the expected signs, were significantly different from zero at the 95% level of confidence, and indicated reasonable elasticities of transaction price with respect to the independent variables. The elasticity of price is greatest with respect to distance to the nearest station and mean income in the area. The value for the coefficient for the parking lot variable was much less than that for the floor area variable, as expected. The CBD versions of the ratio variables were approximately four times greater than their non-CBD counterparts; all ratio variables had positive signs. For the ratio of retail employment density to population density, a positive sign implies that agglomeration effects outweigh the competitive supply effects of retail establishments; the more retail stores per person there are in an area, the higher the transaction price. For the ratio of non-retail employment to retail employment, a positive sign implies that the greater the number of office and other non-retail employees there are in an area per number of retail employees, the more business per retail employee, and the greater the value of the retail store property. These results are all quite plausible.

c) Box and Cox Analysis

Equation 3.5.2 was used in a Box and Cox analysis of the transformation of DPRICE. The estimated value of λ in these runs was 0.1. This value was not significantly from zero at the 90% confidence level, but was significant at 95%.* The estimates for the final Box and Cox model are as follows:

$$(3.5.3)\qquad \frac{DPRICE^{.1} - 1}{.1\ gm(DPRICE)^{-.9}} = \underset{(-.38)}{-\ 15{,}404} - \underset{(-2.02)}{19{,}197}\ \ln(DIST)$$

$$- \underset{(-1.87)}{5{,}671}\ \ln(YR + 1) + \underset{(4.52)}{19{,}104}\ \ln(INCOME)$$

$$+ \underset{(4.19)}{4{,}440}\ \ln(1 + PKLOT) + \underset{(8.35)}{15{,}912}\ \ln(FLAREA)$$

$$- \underset{(-1.57)}{9{,}293}\ (PRE) + \underset{(5.33)}{11{,}319}\ \ln(R/PCBD+1)$$

$$+ \underset{(2.78)}{3{,}028}\ \ln(R/PNCBD+1) + \underset{(4.75)}{9{,}694}\ \ln(QCBD+1)$$

$$+ \underset{(1.78)}{2{,}616}\ \ln(QNCBD+1) + \underset{(2.07)}{15{,}106}\ (HIGHRISE)$$

$$R^2 = .559$$

where the geometric mean of the deflated transaction price was $ 28,319.

While there is a question about the statistical validity of the null hypothesis that λ = 0 (i.e. that a logarithmic transformation of the dependent variable DPRICE is most appropriate), it was decided that the relatively small differences in predictions between the logarithmic model and the Box and Cox results made the simpler, logarithmic form preferable.

---

* The value of the χ^2 statistic with one degree of freedom was 5.22. The critical values 95% and 99% confidence levels are 3.84 and 6.64 respectively.

3.6 - Summary and Evaluation of Model Estimation Results

In the cases of the multi-family dwelling and retail establishment models, it was possible to select one functional form as a final, or "best" model. These results correspond to equations 3.4.2 and 3.5.2 in the preceding sections of this chapter. However, for single family dwellings, no one model was clearly superior. For this reason, three functional forms, equations 3.3.3, 3.3.5 and 3.3.6, were used in later analyses. These models correspond to the best linear additive, mixed log/linear and Box and Cox models respectively.

The model results indicate the following:

(1) In all cases, the distance of a parcel to the nearest Metro station was a statistically significant determinant of the transaction price of an urban parcel. In all of the final models, increasing distance to the station was associated with lower property values; moreover, the effect of distance seems to decline quite rapidly with increasing distance.

(2) The table below summarizes the estimated elasticity of deflated transaction price with respect to distance to the nearest Metro station.

TABLE 3.3 - Elasticities of Price with Respect to Distance to Station

	Single Family			Multi-Family	Retail
Equation No.	3.3.3	3.3.5	3.3.6	3.4.2	3.5.2
Elasticity of Price with respect to Distance	-.13*	-.11*	-.06*	-.19	-.68
Standard Error of Elasticity estimate	.042	.044	**	.0049	.337

* Elasticity evaluated at average of dependent and independent variables.

** The standard errors of the estimated coefficient for the elasticity in the Box and Cox models depends on the variance-covariance matrix of all the estimates. The appropriate calculations would require a separate computational step which was not performed for reasons of time and budget.

As this table indicates, the effect of the Metro system has been far more pronounced in the retail property sector than in either of the residential property markets. Indeed, the elasticity estimated for the retail sector is at least intuitively too high, and should perhaps be viewed as an upper bound.

(3) The effect of the opening date of a particular Metro station on property values is substantial. The table below summarizes the elasticities implied by the model results.

TABLE 3.4 - Elasticities of Price With Respect to Years to Completion

	Single Family		Multi-Family		Retail
Equation No.	3.3.3	3.3.5	3.3.6	3.4.2	3.5.2
Elasticity of Price with respect to number of years to completion	-.11*	-.11*	-.05*	-.09*	-.15*
Standard Error of Elasticity estimate	.050	.038	**	.040	.080

The effect of the number of years to completion appears to be much more uniform over the markets than in the case of transit access. However, the effect is still greatest in the retail property market.

(4) The effects of the other Metro-related variables SS (a dummy variable indicating if the nearest transit station is above ground), PR (a dummy variable indicating if the nearest transit station is a park-ride facility) and PXDUM (a dummy variable indicating extreme proximity to a station) are not certain. None of these variables were included in the final models; careful analysis of the data indicated that in many

---

* See footnote on previous page.

** See footnote on previous page.

cases the data was insufficient to obtain reliable estimates of these coefficients.

(5) Many of the other factors which were hypothesized to affect property values (both parcel-related and demographic variables) appear to have a major influence on both the residential sectors. Included in the demographic set are income, employment densities and the quality of the housing stock; included in the parcel-related variables are the distance to Metro Center and lot area. The racial composition of the neighborhood and a CBD dummy variable were only significant in the multi-family model, and a zoning compatability indicator appeared only in the single family market equation.

(6) The availability of parking (as indicated by the proxy variable PKLOT) positively influences retail property values, though as one might expect, the marginal effect of an added unit of floor space exceeds the effect of an equal amount of parking space. (The elasticities for parking area and floor area are .153 and .562 respectively).

(7) There appear to be strong agglomerative effects on retail property values in the CBD as indicated by a large positive coefficient on R/P CBD (the number of retail employees per person). These effects are much less in non-CBD areas; the estimated elasticities are .399 and .107 for inside and outside the CBD respectively.

(8) The density of non-retail employees per retail employees increases retail properties' values significantly, with a gerater effect within the CBD than outside it. The elasticities of retail properties' values with respect to the number of non-retail employees per retail employees are .342 and 1092 within and outside the CBD respectively.

These conclusions are derived directly from an analysis of the estimated models. In the following chapter, the effects of various factors on specific property values are explored in a series of empirical case studies on specific station areas. These results provide a further basis for drawing inferences from the study.

4. CASE STUDIES OF STATION AREAS

There are three basic purposes in performing a study of specific station areas. These are:

1) to further explore the implications of the models' results, particularly the effects of transit-related variables on property values;
2) to compare the forecasts from the alternative models of the single family dwelling market;
3) to demonstrate how the models might be applied to analyze alternative transit policies (e.g. changes in construction schedules or station location).

In the description of the case studies below, it is important to remember that nearly all of our observed transaction prices are from years during which Metro-rail had not yet begun any of its operations. Consequently, the results of our models reflect how buyers and sellers anticipated the increment in value due to the Metro System. In all probability, models developed using data from ten years hence would produce noticeably different results as a consequence of the system's having been in full operation for several years.

Each case study involves the application of one or more of the models described in Chapter 3 to predict how the values of specific properties are affected by altering some facet of Metro design. The case studies are focused on selected transit stations; in each case, a sample of properties from the station area is used to represent the characteristics of that area.* To the extent to which the subsample of properties in a particular station

---

* These subsamples were also part of the sample used for estimating the models.

area is representative of the station area's characteristics, the forecasts for the station area sample can be considered to be an estimate of the station-area wide effect.

Choosing suitable areas for case studies required the development of criteria. First, we wanted areas for which the conditions were reasonably represented by our model. For example, neighborhoods undergoing major urban renewal or shifts in land use would probably have property values which would not be reliably predicted in our models. Second, we wanted areas which as a set, reflected the diversity of station area conditions in Washington, D.C. Third, it seemed desirable to choose areas which had potential interest to policy makers, in either judging past decisions or formulating future guidelines. For example, the construction of several stations is currently being reconsidered. Fourth, since it was infeasible to obtain new data, the stations in the case studies had o be reasonably well-represented in the overall samples. Finally, the opening of stations in the last eighteen months has raised many questions about the more immediate impact of Metro on property values at these particular stations.

The following section (4.1) is an analysis of the predictions of the three single family dwelling models (equations 3.3.3, 3.3.5, and 3.3.6 in the preceding chapter) for properties in the area of the Potomac Avenue station. This case study provides a comparison of the forecasts yielded by the different models. Section 4.2 extends the Potomac Avenue results to make predictions of the effect of altering particular independent variables to reflect changes in transit system design.

In Section 4.3, the model of the multi-family market is used in analysis of three transit station areas: Brookland, Columbia Heights and Stadium Armory. These three stations represent quite distinct areas and their respective case studies indicate the predictive capabilities of the multi-family model under a range of conditions and the effect of alternative transit designs on property values in different neighborhoods.

Section 4.4 is analogous to Section 4.3 in structure, but uses the model for retail property values on data from properties in the Dupont Circle Station area.

The final section summarizes the major implications of the case studies.

4.1 - Potomac Avenue/Single Family Base Case

Aside from meeting all of the criteria discussed above, the Potomac Avenue station area was chosen because of the stable residential nature of the surrounding neighborhood. In addition, since the Potomac Avenue Metro station was scheduled to be opened in summer, 1977, its opening would be expected to already show a strong anticipatory effect on single family property values in the area.

The Potomac Avenue station area can be characterized as predominantly low to medium-density residential (primarily rowhouses). The major commercial uses in the area are along Pennsylvania Avenue which is strip zoned for commercial uses. Commercial development consists mainly of small retail and service establishments serving the local market. Schools and related playground space are a significant land use in the area.

Present land use policy calls for the preservation of the area for low to medium-density residential uses, with an intensification of development along Pennsylvania Avenue and the area immediately adjacent to the metro station. The D.C. Zoning Commission has adopted as general policy that any re-zoning for the area will be for uses related to neighborhood needs, and will limit initial consideration of zoning changes related to the transit stop to projects at least partially located within 500 feet of the station portals.*

Since the Potomac Avenue station area is a stable residential section, the opening of the station is not expected to have a major impact on development. As such, the area is an ideal place in which to test and compare the single family dwelling models.

A total of 21 observations were used in the Potomac Avenue case study.** Observations differ primarily in the DIST (distance to station), YR (number of years until completion), and LOTA (lot area) variables, whose important statistics are summarized in the following table:

Table 4.1 - Potomac Avenue Sample Statistics

Variable	Mean	Standard Deviation	Range
DPRICE	$24,080	$ 11,817	$ 8,841 - $46,259
DIST	0.25	0.07	0.1 - 0.3
YR	2.14	1.26	0 - 5
LOTA	1,584	559	872 - 2925

Sample Size = 21

* See Municipal Planning Office (1976). "District of Columbia Metro Impact Studies, Progress Report".

** Figures from the 1970 Census indicate the total number of one-unit structures in the Potomac Avenue station area to be approximately 1,590. Those examined comprise about 1% of the total number of single family structures in the station area.

Three functional forms of the single family dwelling model (corresponding to equations 3.3.3, 3.3.5 and 3.3.6) were used in the analysis. These models were for convenience labeled the linear, mixed log/linear and Box/Cox models respectively. The purpose of the initial base case was to compare the three forms as to the estimates they produced. Forecasts using the three model forms produced the following results:

Table 4.2 - Base Case Model Results for Potomac Avenue

	Actual Price	Predicted Price		
		Linear	Mixed Log/Linear	Box/Cox
Mean ($)	24,080	25,247	22,706	23,676
Standard Deviation	11,817	4,328	4,440	3,423

The Box/Cox model produced the mean estimated price closest to the actual price. The percentage difference between the estimated price and the deflated price, defined as

$$\frac{\text{EPRICE} - \text{DPRICE}}{\text{DPRICE}} \times 100\% \qquad \text{where EPRICE} = \text{estimated price}$$

was also calculated for each observation. The mean and standard deviation of the percentage difference for each model form is an indicator of the accuracy of the model. The results are summarized below:

Table 4.3 - Percentage Differences in Model Forms for Potomac Avenue

	Linear	Mixed Log/Linear	Box/Cox
Mean	47%	37%	43%
Standard Deviation	30%	24%	26%

The mixed log/linear form provides the best estimates in terms of minimizing the percentage error between the estimated price and actual price. However, all three forms have relatively large standard deviations, meaning that all three model forms will occasionally have large errors in estimating the transaction prices of individual parcels. All three model forms did not estimate well for prices in the extreme ranges. Percentage errors for observations with extremely high or low prices were, on the average, around 30-40%, even under the best model from.* Generally, the models tended to provide better estimates in the range from $15,000 - $30,000, although some observations within this range were not accurately predicted. The mixed log/linear form predicted prices in the lower end of this range while the linear form was the better estimator at the top of the range. The error pattern for the Box/Cox model appears far more uniform over the range of prices.

*This is not a particularly unexpected result, since implicit in the model formulation is that there is an unobserved disturbance term in the model. Extremely low or high transaction values often correspond to very high or low values of this disturbance, which are not explained in the forecasts.

4.2 - Potomac Avenue/Single Family Case Study

a) delaying the opening of the station

As properties are expected to increase in value as the opening day of a rapid transit stop approaches, a delay in completing such a facility will diminish the amount of value increment that could potentially be recouped at a certain point in time by various value capture mechanisms. One way to "simulate" the effect of a delayed opening is to increase the YR variable (years until completion) for each observation by a certain amount. The results of the policy runs for a delay of one and three years are shown below.

Table 4.4 - Potomac Avenue: Effect of Delay in Station Opening

Policy \ Forecast	Actual Price	Predicted Price		
		Linear	Mixed Log/Linear	Box/Cox
Base Case	24,080*	25,247	22,706	23,671
1 Year Delay	-	24,371 (-3.5%)**	21,321 (-6.1%)	22,130 (-6.5%)
3 Year Delay	-	22,619 (-10.4%)	19,522 (-14.0%)	20,636 (-12.8%)

*mean of case study sample

**percentage change from base case shown in parentheses

The models predict a decrease in value of $875 to $1,540 for a delay of one year. This decrease comprises a 3.5% to 6.5% change in value from the base case.* The mixed log/linear form and the Box/Cox form of the model forecasts seem more plausible than these of the linear model since the anticipator effects of a rapid transit station's opening is strongest when the opening is imminent, a small delay would be expected to have a relatively larger effect on value than a long delay. A long delay would inhibit some of the anticipatory expectations, thereby lessening sharp changes in value.

b) relocation of the Metro Station

Suppose, for example, that the Potomac Avenue station is relocated to be at a natural boundary of the neighborhood rather than in the middle of it, or the station is relocated to be on a large, vacant or underdeveloped tract of land. In both cases, the average single family parcel is now a further distance away from the transit station.**

The results of the policy runs for increments of 0.2, 0.4, 0.6, 0.8, and 1 mile are summarized in Table 4.5 below.

---

*Longer delays result in constant value decrease per year under the linear form of the model. With the mixed log/linear form and the Box/Cox form of the model, longer delays result in diminishing value decreases per year. The Box/Cox form shows the most rapid diminishing of value decreases per year.

**Great care must be taken in the interpretation of this relocation scenario. If the station is relocated to another point in the middle of the neighborhood, the distance from the average parcel may not change. Some parcels will be closer but other parcels will be farther away. The net effect on property values might well be zero or even positive.

Table 4.5 - Potomac Avenue: Effect of Relocation of Station

Type of Forecast / Policy	Actual Price	Predicted Price		
		Linear	Mixed Log/Linear	Box/Cox
Base Case	24,080	25,247	22,247	23,671
DIST + .2	-	22,371 (-11.4%)	20,330 (-10.5%)	21,116 (-10.8%)
DIST + .4	-	21,471 (-15.0%)	19,587 (-13.7%)	20,482 (-13.5%)
DIST + .6	-	21,014 (-16.8%)	19,209 (-15.4%)	20,107 (-15.1%)
DIST + .8	-	20,735 (-17.9%)	18,979 (-16.4%)	19,897 (-15.9%)
DIST + 1.0	-	20,544 (-18.6%)	18,939 (-16.6%)	19,749 (-16.6%)

The results of these policy runs seem sensible in that the relocation of a station a short distance away from the average parcel should have a significant effect as the station becomes less accessible by foot. Since the average distance for the case study sample was 0.25 miles, an increment of 0.2 miles results in an average distance of 0.45 miles. A single family dwelling this distance away from a transit station is still within easy walking distance, although it is not as ideal as if it were closer. Consequently, the value of a dwelling, other things being equal, should decline substantially because of this decrease in pedestrian accessibility. A further increment of 0.2 miles results in the dwelling being 0.65 miles away. At this distance, and beyond, the dwelling is not within easy pedestrian access to the station. Consequently, property values should not be greatly influenced by the station accessibility, as shown by the rapidly diminishing percentage change in value changes.

c) eliminating a station along a line

The Potomac Avenue station is 0.6 miles from the Eastern Market station and 0.7 miles from the Stadium/Armory station. Under this scenario, the Potomac Avenue station is eliminated and the Metro line will run 1.3 miles non-stop from the Eastern Market station to the Stadium/Armory station. Consequently, the closest Metro station will be either the Eastern Market or Stadium/Armory station.

All of the observations in the case study sample were found to be closer in distance to the Eastern Market station. Consequently, they were all assigned to that station. The technique used to simulate the elimination of the Potomac Avenue was to assign the YR variable (years until completion) for each observation to that of the Eastern Market station. In addition, the DIST variable (distance to nearest station) for each observation was replaced by the new distance to the Eastern Market station.

The results of the policy runs are summarized below.

Table 4.6 - Potomac Avenue: Effect of Elimination of Station

Type of Forecast / Policy	Actual Price	Predicted Price		
		Linear	Mixed Log/Linear	Box/Cox
Base Case	24,080	25,247	22,706	23,671
No Station	-	22,525 (-10.8%)	20,457 (-9.9%)	21,322 (-9.9%)

All three forms of the model predict that the average decline in single family property value will be approximately 10% of the original estimated base value.

4.3 - Multi-family Building Case Studies

a) station area descriptions

Brookland, Columbia Heights and Stadium/Armory stations were chosen for case studies using the final, multi-family building model (equation 3.4.2). Brief profiles of each of these station areas (as we have defined them) can be found in Table 4.7. The area around the Brookland station is primarily residential in character, with single family units most prevalent. Nearly half of all units are owner-occupied and mean income, $11,600, is well over the District's average of $9,600. Relative to the entire city, the percent non-white is below average (60% versus 73% in 1970). Unlike the Columbia Heights station, there is no question whether the Brookland station will be opened; as of February, 1978, its area's residents enjoy Metro-rail service. Columbia Heights also contrasts Brookland in that only 18% of its residents own the units in which they live. The station area is predominantly non-white and is one of the District's poorest and highest crime sections, and has one of the lowest auto ownership levels. Since the station is now part of an UMTA-mandated "alternatives analysis," its exact location is still not precisely fixed. The Stadium/Armory station fits somewhere between these two stations. The surrounding area is very stable, residential in character, has 40% owner-occupied dwellings and almost matches the District's average income level. Its residents have had Metro-rail service since early July, 1977.

Table 4.7 - Profile of Multi-family Case Studies' Station Areas*

Attribute / Station Area Name	Distance from Metro Center	Percent Owner-occupied Dwellings	Percent Non-white	Mean Income	Station Status
Brookland	3.7 miles	45%	60%	$11,600	opened Nov. 77
Columbia Heights	2.7 miles	18%	93%	$7,900	in question
Stadium/Armory	3.8 miles	40%	95%	$9,700	opened 3 July 1977

*all demographic data as of 1970

Table 4.8 summarizes the samples used in the case studies for the three station areas. Included are the means of the actual, deflated transaction price (DPRICE), the distance to nearest Metro station (DIST), the number of years between the transaction and the station's scheduled opening (YR), the lot area (LOTA) and the number of dwelling units (NBDU).

Table 4.8 - Sample Characteristics for Multi-family Building Case Studies

Attribute / Station	Sample Size	DPRICE	DIST	YR	LOTA	NBDU
Brookland	40	$25,370	.45	3.05	3,574	6.06
Columbia Heights	40	$51,390	.34	6.18	5,393	14.2
Stadium/Armory	70	$26,152	.55	2.65	3,457	7.20

As in the Potomac Avenue case study, a base case analysis was performed to test how well the models predicted transaction prices in each station area. The results of these tests are given in Table 4.9 below. The magnitude of the percentage errors for the multi-family model are roughly comparable to those for the single family models.

Table 4.9 - Base Case for Multifamily Dwelling

Station Area	Actual Mean Price	Predicted Mean Price	Mean Percentage Difference Between Estimated and Actual Price*
Brookland	$25,370	$28,378	30%
Columbia Heights	$51,390	$45,520	34%
Stadium Armory	$26,152	$28,155	39%

*Defined as in section 4.1

b) delaying the opening-up of the stations

The same type of policy testing performed in the Potomac Avenue case study was done for the three multi-family case studies. The results of these tests for changes in the number of years until the station is opened by one and three years are summarized in Table 4.10. As expected from the functional form of the model, for the two stations which have already opened, (Stadium / Armory and Brookland) increasing the "number of years to completion" by two and then by four years made more of a relative difference than in the area whose station was, one the average, over six years from being completed.

Table 4.10 -Multi- family Building Case Studies: Delay in Station Opening

Policy / Station Area	BASE		1 Year Delay		3 Year Delay	
	Actual Price	Predicted Price	Predicted Price	Percentage Change from Base Case Forecast	Predicted Price	Percentage Change from Base Case Forecast
Brookland	$25,370	$28,378	$27,663	-2.6%	$26,580	-6.2%
Columbia Heights	$51,390	$45,520	$44,857	-1,5%	$43,770	-3.8%
Stadium/Armory	$26,152	$28,115	$27,254	-2.9%	$26,102	-6.9%

c) <u>decreasing the number of stations</u>

Table 4.11 presents the model forecasts when the value of the distance to the nearest station variable (DIST) is altered to reflect a decrease in the number of stations along the transit lines. The predicted impacts tend to be relatively small, particularly for shifts which produce changes in access of .2 miles. Similar predictions for the single family dwellings produced percentage changes in value which were on order of twice as great for small shifts in DIST, but of comparable magnitude for large changes (e.g. a mile increase).

Table 4.11 - Multi-family Building Case Studies: Relocation of Station

Policies / Station Area	Actual Base	PREDICTED MEAN PRICES AND PERCENT CHANGES*					
		Predicted for Base	DIST+.2	DIST+.4	DIST+.6	DIST+.8	DIST+1.0
Brookland	$25,370 (-)	$28,378 (-)	$26,725 (-5.7%)	$25,548 (-9.8%)	$24,642 (-13.0)	$23,910 (-15.0%)	$23,298 (-17.7%)
Columbia Heights	$51,390 (-)	$45,520 (-)	$42,219 (-7.0%)	$40,044 (-11.6%)	$38,438 (-15.1%)	$37,173 (-17.9%)	$36,135 (-20.1%)
Stadium/ Armory	$26,152 (-)	$28,155 (-)	$26,432 -5.6%	$25,254 (-9.5%)	$24,360 (-12.5%)	$23,642 (-14.9%)	$23,043 (-17.0%)

* Percent changes given in parentheses.

4.4 - Dupont Circle/Retail Establishment Case Study

a) area description

Dupont Circle was chosen as a case study area for the retail establishment model for several reasons. First, the Dupont Circle station had already opened, and hence, actual value changes due to improvement in transit accessibility could be observed. Second, although its mean income was higher than average ($12,900 compared to $9,600 for the District) and percent non-white lower than average (25% compared to 63% for the District), the area had been relatively stable in terms of neighborhood quality and retail activity*. A sample of 24 observations were used in this case study. The subsample means of the most important variables in the retail establishment model are given in Table 4.12.

Table 4.12 - Dupont Circle Sample Characteristics

Variable	Means
DPRICE ($)	45,137.20
DIST (mi)	0.26
YR (yrs)	1.2
RENTER (%)	87
INCOME ($/yr)	12,914
DISTM (mi.)	1.3
LOTA (sq.ft.)	1,489.87
FLAREA (sq.ft.)	1,547.17
EDEN (employees/sq.mi.)	52.889.2
RDEN (employees/sq.mi.)	4,154.33
PDEN (persons/sq.mi.)	43,321.5
Sample Size = 24	

*Municipal Planning Office (1976): "Ward Two: Summary Policy", District of Columbia.

b) base case

As indicated in Table 4.12, the base year mean deflated price of the 24 properties in the Dupont Circle Station area was $45,137. The average of the predicted prices for the same properties was $40,965. The average percent error in the individual perdictions was 63%.

c) delaying the opening of the station

By increasing the YR variable and comparing the resulting mean transaction price with the base case, the relative effect of delaying the opening of a station can be observed. Increments of 1 and 3 years were added to completion variable and the changes in the mean predicted transaction prices were as follows:

Table 4.13 - Dupont Circle: Effect of Delay in Station Opening

Policy / Base Case	Actual Price	Predicted Price	Percent Change from Predicted Base
Base Case	$45,137	$40,965	-
1 Year Delay	-	$28,265	-31 %
3 Year Delay	-	$25,398	-38 %

The model predicts that if a station is definitely to be built, the delay of an additional year would result in a 31% decrease in what the property would be worth if there were no delay. These predicted changes seem, at least on an intuitive basis, somewhat too high.* It would appear that the model is overly sensitive to the years until completion variable, and that the above figures should probably be taken as reasonable upper bounds.

d) decreasing the number of stations

The decision to decrease the number of stations in the systems can be represented by increasing the value of the distance to station variable, DIST. In other words, the average distance of retail properties to a station will increase if fewer stations are built. As in the previously described case studies, increments of 0.2 mile were added to the observed distance, and the results of the policy runs were as follows:

Table 4.14 - Dupont Circle: Effect of Relocation of Station

Policy \ Type of Forecast	Actual	Predicted	Percent Chage from Predicted Base
Base Case	$45,137	$40,965	-
DIST + .2	-	$27,856	-32%
DIST + .4	-	$25,398	-38%
DIST + .6	-	$23,350	-43%
DIST + .8	-	$22,121	-46%
DIST + 1.0	-	$20,483	-50%

The results imply that if a station area was anticipated, and if the elimination of that station area from the system would increase the average distance to the next nearest station by 0.2 mile, there would be a 32% decrease in the average transaction price of properties in the original station area. Because of the functional form of the model, additional increments in average distance

* This result is not surprising considering the high value of the estimated elasticity of deflated price with respect to distance.

to the nearest station have decreasing marginal effects on transaction price. As with the previous results, the effect of the first small change seems higher than one would usually anticipate. The marginal effects beyond the first .2 miles seem much more reasonable than the shift of 32% for the first distance increment. The larger than expected shift with changing distance is a reflection of the very high estimated elasticity of price with respect to distance, and should probably be viewed as a "high end" estimate.

4.5 - Summary of Case Studies

The case studies described in this chapter are obviously quite limited and any conclusions drawn from them should be treated as tentative at best. Given this caveat the case studies appear to indicate the following:

1) The models have a limited ability to predict the prices of individual transactions accurately; there are simply too many factors which are not appropriately measured by the available data. However, for samples of parcels from a specific geographic area, (i.e. areas around transit stations), there was no clear under- or over-estimation of transaction prices. Thus, the models appear to be useful for predicting average property value within limited geographical areas even though they were estimated using data from all of Washington, D.C.

2) Of the three single family dwelling models (linear, mixed log/linear and Box/Cox), the Box/Cox model best replicated the mean of the Dupont Circle properties in our sample. Moreover, the Box/Cox model seemed to be more useful over a wide range of parcel types than the other two forms.

3) Retail transaction prices appear to be more sensitive than single or multifamily prides to delays in station openings and increases in access distance. The forecasts from the retail model, however, appear to be unreasonably high and probably should be used as an upper bound in any analysis.

4) The most basic components of transit system design such as section opening schedules and station spacing appear to have influences which are not only statistically significant (as indicated in Chapter 3) but also numerically important in an absolute sense.

5. CONCLUSIONS

As discussed in Chapter 1, the concept of using public action to "tax away" or otherwise share some of the benefits of public investment in urban transit investments is based on a number of fundamental assumptions, the most significant of which is that such benefits are capitalized in the real estate market. This study has provided some empirical support for the possibility that real estate property shifts do indeed occur in areas near transit stations. However, it leaves open the other questions about the economic efficiency and equity associated with alternative "value capture" policies.

It is clear that the decision of whether or not a value capture program will be implemented in conjunction with mass transit investments is largely a political one. Issues of equity among various groups (particularly land owners and the public-at-large) as well as the need for increased revenue to offset the escalating costs of transit construction and operation are likely to dominate in the decision-making process. Nevertheless, research of the type described in this report can help pinpoint areas of maximal potential or that combination of transit-related, parcel-specific and demographic features which lead to the greatest increase in values.

For example, as indicated in the conclusion of Chapter 3, the value of retail properties appears to be much more sensitive to proximity to transit stations. The relatively high elasticity of retail property value with respect to the distance to the nearest Metro station suggests that retail areas are better suited for any form of value capture policy ranging from direct taxation to joint private/public sector development of retail floor space.

Although our research has broken new ground in several respects, there have also been a number of issues raised, both explicitly and implicitly, which could not be resolved within the time and budget limitations of this project. Some of the most critical of these areas are as follows:

1) It is still necessary to examine transaction prices in Washington, D.C. for the period after Metro was opened. There is no certainty that the kind of anticipatory reaction found prior to the system's opening would continue. Even around stations which will not open for a number of years (if at all), the anticipatory reaction may be different simply because Metro is a reality for other neighborhoods. Future research might monitor the development of the value of properties for the annual periods following 1976. This evidence would provide the only clear picture of how urban property values near transit stations evolve through the planning, constructing, and operating phases of a transit system.

2) To the extent that the real estate market in Washington, D.C. is unique, it would also be useful to conduct research parallel to the present effort in other cities which anticipate major mass transit investments. Properties in cities with real estate markets which are less active than Washington's would probably show far slower rates of change and may have smaller transit-related impacts.

3) In conjunction with (2), it is important to consider the variations in impact on real estate markets of different types of transit systems. That is, does the presence of a light rail system or a system of exclusive bus lanes have substantially different impact on the values of urban properties than does a heavy rail system as in Washington, D.C.? The answer to such questions would prove invaluable in the evaluation of alternative transit system designs.

4) It is apparent that virtually any major physical investment can have an impact on the real estate market. Consequently, the impact of transit in conjunction with other major investments should be examined. Housing developments, commercial and retail space, recreational areas, and parking facilities have all been developed on or near transit stops. This kind of "joint development" is surely a means to encourage values to increase. Unfortunately, it is not at all clear how much effect these facilities have and the extent to which transit-related effects interact with these investments to induce value shifts.

5) Given adequate funding, one could supplement our data base with data which was either too costly or too time consuming to obtain. The most significant of these effects are rent levels, value of improvements in properties and structural condition.

6) Our study has been restricted to changes in property value; in reality, transportation systems may also alter patterns of property use. Our limited empirical evidence suggests, for example, that values will rise faster for retail property than residential property. This would imply strong economic incentives for some conversion of land use near transit stations, which, having zoning restrictions, may impact on specific segments of the housing market. Some examinations of whether this is indeed occurring is clearly warranted.

7) The effect of transit investment on vacant parcels was beyond the scope of this study. The basic methodology of our study, however, is applicable to such parcels and would provide significant insights into the effects of a transit system in relatively underdeveloped areas.

8) Finally, as systems like Metro expand, it will become more important to extend the focus of our study to suburban portions of the Metropolitan area.

The value of the above research would not be restricted to improvements in policy designed to increase the proportion of transit-induced property value changes accruing to the public sector. Transportation and other public infrastructure investments play a vital role in shaping urban form, and to a great extent, changes in property values are an indicator of how public action is evaluated in the marketplace. An adequate understanding of how such values respond to transit decisions, provide perhaps one of the clearest (but by no means the only) mechanism for evaluating the benefits associated with such systems.

REFERENCES

1. Quantitative Analyses of the Determinants of Property Values

Adkins, William G. Effects of the Dallas Central Expressway on Land Values and Land Use. Texas Transportation Institute, a part of the Texas A & M College system. Bulletin #6. 1958.

Adkins, William G. "Land Value Impacts of Expressways in Dallas,Houston, and San Antonio, Texas." Highways and Economic Development. Highway Research Board. Bulletin 227. 1959.

Adkins, William G., and Alton W. Tieken. "Economic Effects of Expressways Without Paralleling Service Roads in the City of San Antonio." Texas Engineering Experiment Station News. March 1959.

Allen, Bruce W. and Robert Mudge. The Impact of Rapid Transit on Urban Development: The Case of the Philadelphia-Lindenwold High Speed Line. Rand Corporation Paper P-5246. 1974.

Ashley, R. N., and W. F. Bernard. "Interchange Development Along 180 Miles of I-94." Highway Research Record, No. 96. Highway Research Board. 1965.

Boyce, David, Bruce Allen, Richard R. Mudge, Paul B. Slater and Andrew M. Isserman. The Impact of Rapid Transit on Suburban Residential Property Values and Land Development Analysis of the Philadelphia High Speed Line. Regional Science Department, Wharton School, University of Pennsylvania. NTIS PB 220 693. November 1972.

Burton, Robert C., and Frederick D. Knapp. "Socio-Economic Change in the Vicinity of the Capital Beltway in Virginia." Highway Research Record, No. 75. Highway Research Board. 1959.

Clonts, Howard A., Jr. "Influence of Urbanization on Land Values at the Urban Periphery." Land Economics. November 1970. pp. 489-496.

Cribbins, P.D., W.T. Hill, and H.O. Seagraves. "Economic Impact of Selected Sections of Interstate Routes on Land Value and Use." Highway Research Record, No. 75. Highway Research Board. 1959.

Czamanski, S. "Effects of Public Investments on Urban Land Values." Journal of the American Institute of Planners. July 1966.

Davies, Gordon W. "The Effect of a Subway on the Spatial Distribution of Population." Journal of Transport Economics and Policy. 1972. pp. 126-136.

Davies, Frederick W. "Proximity to a Rapid Transit Station as a Factor in Residential Property Values." The Appraisal Journal. October 1970. pp. 554-572.

Downing, Paul B. "Factors Affecting Commercial Land Values: An Empirical Study of Milwaukee, Wisconsin." Land Economics. 1973. pp. 44-56.

Eyerly, R.W. Property Formation, Real Estate Value, and Land Use in Four York County Interchange Communities: An Interim Report. Pennsylvania State University. 1965.

Goldberg, Michael A. "Economics of Transportation Corridors: Further Empirical Analysis." Highway Research Record, No. 410 1972. pp 37-51.

Golden, J.S. "Land Values in Chicago: Before and After Expressway Construction." Chicago Transportation Study. October 1968.

Langfeld, S.C. The Balanced and Orderly Development of the Site in Close Proximity to a Metro Station as a Contributor to a More Healthy Economically Viable Urban Environment in the Washington Metropolitan Area. Urban Transportation Center. NTIS PB 203783. 1971.

Lee, Douglas B., Jr. et al. Case Studies and Impacts of BART on Prices of Single Family Residences. Volumes III and VI of Land Use and Investment Study. Part II of BART II: Pre-BART Studies of Environment, Land Use, Retail Sales. Berkeley: Institute of Urban and Regional Development of the University of California. Prepared for the Metropolitan Transportation Commission. 1973.

Lemly, J.H. "Changes in Land Use and Value Along Atlanta's Expressways." Highways and Economic Development. Highway Research Board Bulletin 227. 1959.

Miller, S.G. Effects of Proposed Highway Improvements on Property Values. National Cooperative Highway Research Program Report No. 114. Highway Research Board, 1971.

Pendleton, William C. "Relation of Highway Accesibility to Urban Real Estate Values." Highway Research Record, No. 16. 1963. pp.14-23.

Spengler, E.H. Land Values in New York in Relation to Transit Facilities. New York: Columbia University Press. 1930.

2. Highway and Transit Systems as Tools in Urban Planning

Anderson, Arnold C. "The Effect of Rapid Transit on Property Values." The Appraisal Journal. January 1970. pp. 59-68.

Averous, Christian P., and Douglas B. Lee, Jr. "Land Allocation and Transportation Pricing in a Mixed Urban Economy." Journal of Regional Science, 18, 2. 1973.

Barden, R., and J.H. Thompson. The Urban Frontier: Occasional Paper No. 4. Syracuse University, Urban Transportation Institute. NTIS PB 204 056. 1971.

Boston, Massachusetts, Metropolitan Area Planning Council. Joint Planning Guidelines. Richard M. Doherty, Executive Director. March 1975.

Burkhardt, Ross. Summary of Research: Joint Development Study. The Administration and Managerial Research Association of New York City. Mayor's Midtown Action Office, 42nd St., 23rd Fl., New York City, 10036. 1976.

Carroll, Donald D., John R. Borchert, James Schwinden, Philip M. Raup. The Economic Impact of Highway Development Upon Land Use and Value. University of Minnesota. Study submitted to the Bureau of Public Roads, U.S. Department of Commerce. 1958.

Corsi, Thomas M. "A Multivariate Analysis of Land Use Change: Ohio Turnpike Interchanges." Land Economics. August 1974. pp. 232-241.

Davis, Frederick W. "Proximity to a Rapid Transit Station as a Factor in Residential Property Values." The Appraisal Journal. October 1970. pp. 554-572.

Gauthier, H.L. "Geography, Transportation, and Regional Development." Economic Geography. October 1970.

Goldberg, Michael A. "Transportation, Urban Land Values, and Rents: A Synthesis." Land Economics. May 1970. pp. 153-162.

Hamburg, John R., and Geoffrey J.H. "Impact of Transportation Facilities on Land Development." Paper sponsored by the Committee on Land Use Evaluation and present at the 49th Annual Meeting of the Highway Research Board. Highway Research Record, No. 305. 1970.

Heenan, G. W. "The Economic Effect of Rapid Transit on Real Estate Development." Appraisal Journal. April 1968.

Kentucky, University of. The Effect of the Louisville-Watterson Expressway on Land Use and Land Values. Lexington, Kentucky: University of Kentucky Press. 1975.

Lee Douglas B., Jr., and C. P. Averous. "Land Use and Transportation: Basic Theory." Environment and Planning. 1973. pp. 491-502.

Mohring, Herbert. "Land Values and the Measurement of Highway Benefits." Journal of Political Economy. 79. 1961. pp. 236-249.

Soot, Siim. "Transportation Costs and Urban Land Rent Theory: The Milwaukee Example: 1949-1969." Land Economics. 1973.

Waldo, Robert D. "Urban Land: Values and Accessibility." Land Economics, Reports and Comments. 1974. pp. 196-201.

3. Analyses Which Discuss Methodological Issues

Box, G.E.P., and D.R. Cox. "An Analysis of Transformations." Journal of the Royal Statistical Society, Series B. 1964. pp. 211-243.

Beesley, M.E., and C.D. Foster. "Estimating the Social Benefit of Constructing an Underground Railway in London." Urban Transport. London: Butterworths. 1973. pp. 1-97.

Boyce, David E., and Bruce Allen. Methods and Strategies for Assessing Transportation Impacts on Land Use and Urban Development: BART Working Paper No. 1. BART Impact Program Paper 3. September 1976.

Boyce, David, and Bruce Allen. Research Recommendations, BART Impact Program Working Paper, No. 2. 1974.

Brigham, E.F. A Model of Residential Land Values. The Rand Corporation. Santa Monica, California. 1964.

Brodsky, Harold. "Residential Land and Improvement Values in a Central City." Land Economics. August 1970.

Clonts, Howard A., Jr. "Influence of Urbanization on Land Values at the Urban Periphery." Land Economics. November 1970. pp. 489-496.

Dornbusch, David M. "BART-Induced Changes in Property Values and Rents." Land Use and Urban Development Project, Phase I, BART Working Papers. WP 21-5-76. U.S. Department of Transportation, U.S. Housing and Urban Development. 1976.

Hammer, Thomas R., et. al. The Effect of a Large Urban Park on Real Estate Value. Regional Science Research Institute, Discussion Paper Series: No. 51. September 1971.

Isibor, E. I. Modelling the Impact of Highway Improvements on the Value of Adjacent Land Parcels. Purdue University, Indiana State Highway Commission. 1975.

Lee, Douglas B., Jr. "Analysis of BART Impacts on Bay Area Land Use." Transportation Engineering Journal. Proceedings of the American Society of Civil Engineers. Institute of Urban and Regional Development, University of California, Berkeley. Reprint No. 78. May 1972.

Onibokum, Adepoju G. Socio-Economic Impact of Highways and Commuter Rail Systems on Land Use and Activity Patterns: An Annotated Bibliography. Council of Planning Librarians Exchange Bibliography. #815. June 1975.

Rao, Potluri, and Roger LeRoy Miller. Applied Econometrics. Wadsworth Publishing Co., Inc. Belmont, California. 1971.

Richardson, Harry W., Joan Vipond, and Robert A. Furbey. "Determinants of Urban House Prices." Urban Studies. 11. 1974. pp. 189-199.

Rogers, Andrei, et. al. The Time Lag of Factors Influencing Land Development. Institute for Research in Social Science, University of North Carolina, Chapel Hill, N.C. October 1963.

Seyfried, Warren R. "The Centrality of Urban Land Values." Land Economics. August 1963. pp. 275-284.

Sharpe, Carl P., et. al. A Value Capture Policy: Financial Element. Volume IV. Rice Center for Community Design and Research. Rice University, Houston, Texas. November 1, 1974.

Wilkinson, R. K., and Catherine A. Archer. "Measuring the Determinants of Relative House Prices." Environment and Planning. 5. 1973. pp. 357-367.

Wootan, D. V. and C. R. Haning. Changes in Land Value, Land Use, and Business Activity Along a Section of the Interstate Highway System in Austin, Texas. Texas Transportation Institute, Texas A & M College System, Texas. 1960.

4. Data Sources

District of Columbia. "Ward Summary Profiles." Municipal Planning Office. February 1976.

Lusk's District of Columbia Real Estate Directory Service, Recorded Transfers and Mortgages: January 1966-April 1972 and May 1972-June 1976. Rufus S. Lusk and Son, Inc., Publishers.

"METRO Means Development Boom." Washington Star-News. June 23, 1974.

U.S. Bureau of the Census. 1970 Census, Tract and Block Statistics. Washington D.C. Government Printing Office.

U.S. Bureau of Labor Statistics. Handbook of Labor Statistics, 1975. U.S. Department of Labor, Bulletin 1905. Washington, D.C., Government Printing Office.

"Waiting for the Impact." Mass Transit. July/August 1977. pp. 18-23.

Washington Council of Governments (COG). Metropolitan Parcel File, 1972.

APPENDIX A

A REVIEW AND ANALYSIS
OF THE LITERAURE RELATED TO ASSESSING
THE IMPACT OF TRANSPORTATION IMPROVEMENTS
ON URBAN PROPERTY VALUES

A.I. Purpose, Limitations and Structure of the Literature Review

1. Purpose of Review
2. Limitations of Review
3. Structure of Review

A.II. Elements of Analysis

1. Scope and Characteristics of Study Area
2. Appropriate Dependent Variable
3. Appropriate Independent Variables
4. Data Sources
5. Sampling Procedures
6. Appropriate Analytic Methods
7. Conclusions About Elements of Analysis

A.III. Hypotheses Generated From the Literature

A.IV. Summary of Conclusions

A.V. List of Areas Which Have Been Studied

A.VI. Table of Literature Abstracts

A.VII. Annotated Bibliography

A.I. PURPOSE, LIMITATIONS AND STRUCTURE OF THE LITERATURE REVIEW

A.I.1 - Purpose of Review

The first phase of our study attempted to examine previous work with the objectives of determining the extent to which any of the above questions has been satisfactorily answered. More specifically, the literature search was undertaken for the following six purposes:

(1) To familiarize the project team with the existing body of theories, empirical evidence, and procedures for collecting that evidence,

(2) To review the various methodological approaches which have been used in the past to determine transportation accessibility impacts on property values,

(3) To glean from the literature several hypotheses which could be tested in the modeling effort,

(4) To develop a relevant collection of abstracts from the literature (i.e., an annotated bibliography) for convenient reference,

(5) To tabulate the abstracts of prior work, and

(6) To list the station, interchange, metropolitan, or regional areas which have been previously studies.

Two excellent references which deal with several of the above items are the Joint Development Study, undertaken by Ross Burkhardt and the Administration and Management Research Association of New York City, Inc. (AMRA, 1976) and the Socio-Economic Impact of Highways and Commuter Rail Systems on Land Use and Activity Patterns--An Annotated Bibliography, written by Onibokum (1969) and published by the Council of Planning Librarians. They are, however, limited in the following respects: the relatively new AMRA study emphasizes the circumstances and the degree to which land value changes occur, and the relatively older Council of Planning Librarians Exchange Bibliography emphasizes impacts on land use, and not value.

We attempt not only to combine and sift relevant items from these two efforts, but more importantly, to supplement them with a discussion about current methodology applied to the measurement of transportation accessibility impacts on land value.

A.I.2 - Limitations of Review

The scope of the review is strictly limited to the purposes outlined above. Thus, it will not deal directly with land use impacts, sociological impacts, or measurement of benefits and costs of transportation systems, nor will it pursue the topic of value capture itself in great depth. The emphasis is on study methodology.

The literature review was not intended to be an exhaustive compilation of all relevant work. Rather, the literature search upon which it is based was an explicit attempt to sample a wide range of prior studies and was heavily oriented towards more recent work. The review places heavy emphasis on evaluating the methodological choices that were required in our own study, and places significantly less emphasis on actual quantitative results of the policy implications of the work.

As in all such reviews, there is a clear tradeoff between simply stating what previous reserachers have done and including a critical assessment of the validity and relevance of prior studies. In general, this review represents a middle ground; an effort is made to first report the content of relevant studies and then present our study team's assessment of the merits of the work.

A.I. 3 - Structure of Review

The following section of this document, Elements of Analysis, examines each of six methodological areas in which some decision must be made when designing an approach to measure transit-induced land value changes. Section III, Hypotheses Generated from the Literature, is a listing of specific hypotheses about land value change which have been discussed by previous researchers. Section IV is a summary conclusions derived from the review.

The three remaining sections of this appendix are a table listing the geographical areas prior researchers have studies (A.V), a table of literature abstracts (A.VI), and an annotated bibliography (A.VII).

A.II. ELEMENTS OF ANALYSIS

There are a number of key methodological decisions which must be made in the deveolpment of analysis, these decisions include defining the scope and desired characteristics of the study area, choosing the appropriate dependent and independent variables, data sources, sampling procedures and, of course, the appropriate analytical methodology. Careful planning and use of these elements of analysis are critical to a credible and valid study.

While there does exist in the literature a few studies which describe some of the methodological decisions which they had to make, discussion of most of these elements is sorely lacking in most of the existing empirical literature. Most studies simply state their elements of analysis with little or no discussion about how they arrived at their final decisions on which independent variables to use, which sampling technique to use, etc. This is quite frustrating in an area where very little has been understood or agreed upon. For example, while some analysts feel that choosing independent variables for this type of study is a process which necessitates complicated (and mechanistic) stepwise linear regression or factor analytic methods, others feel that it necessitates only a little intuition. In the midst of this controversy, transportation impact studies, with few exceptions, simply list the variables used with no justification or explanation at all for their inclusion in regression models.

The following sections briefly discuss the major elements of analysis and review the literature with respect to these elements.

A.II.1 - Scope and Characteristics of Study Area

The scope of the study area used in previous studies is naturally constrained to the region surrounding transportation facilities. In addition, the following considerations should be taken into account:

(1) Governmental jurisdictions may not coincide with the impacted region, thus creating data collection problems,

(2) The study area should be compact for cost-efficiency reasons,

(3) The study area should include a wide range of environmental conditions so as to maximize variation in the available data,

(4) The study area should be representative of the different types of impact regions. The criteria for selecting representative regions ought to include housing and commercial types and intensities, as well as different mixes of resident social and economic characteristics (Dornbusch (1975)),

(5) Most of the transit station or limited-access freeway interchange studies used a radial or concentric study area surrounding the access point, while free-access roadway studies used parallel bands surrounding the roadway. Thus, the study area should be chosen with regard to the type and degree of access to the transportation improvement,

(6) The availability of property value information should be investigated (Dornbusch (1975)),

(7) The highway studies which used "bands" paralleling the highway (e.g. Adkins (1958)) considered land outside the outer bands as control areas,

(8) In high-density regions, impacted areas may overlap with one another, thus complicating the analysis,

(9) Negative impacts of the transportation investment may be observed along the transportation alignment.

A. II.2 - Appropriate Dependent Variable

Several of the study efforts in the past showed great concern for the separation of analysis of the land value impact of transportation investments from the impact on the sum of land and improvements values (see Brigham (1964), Hammer (1971)); many others simply glossed over the issue and remained ambiguous as to which they used.

It is necessary to choose a dependent variable which reflects the objective of each particular study. In our case, we are attempting to determine the effects of improved transit accessibility on the values of several different types of land use; therefore, we are using the sum of land and improvements values as the dependent variable. In other studies, the object might be to determine the effects of an influence simply on land, in which case only land values should be used for the dependent variable. (Meaningful data on values for land only, however, are frequently very difficult to obtain for properties which have improvements.) Thus, the issue here really is to clearly define what it is the particular study is attempting to observe. The matter may be complicated by poor data sources (for example, tax assessors' data may not separate land from improvements), but such complications should not obscure the reasoning behind choosing a particular dependent variable.

Given a choice of which value to measure, there is the additional problem of choosing the appropriate measure of property value. The measures which have received the most attention are: sales price, tax assessment value, and appraised value.

There are several potentially conflicting issues involved with choosing the appropriate measure of property value:

(1) The measure should be available for all the relevant time periods. In particular, there may be a problem with tax assessment values; many jurisdictions assess properties in successive years, resulting in a lack of up-to-date data for certain parcels during certain years. There may also be a problem with sales price data; turnover rates in some areas may simply be very low, also resulting in a lack of up-to-date data for certain parcels.

(2) The criteria used to determine the values should be uniform; For example, appraised value of all single family residential properties in the sample should be determined by using the same valuation criteria.

(3) Time lag effects should be minimized. A measure should be chosen such that it would accurately represent any influence of the improved transportation accessibility as soon as the influence is felt.

(4) The measure should be relevant to the goals of the study effort. Thus, if the study is attempting to determine the effect of improved transit access on tax revenues, tax assessment value should be used.

It may not be possible to satisfy each and every one of these conditions with the choice of a particular measure, but they should be considered and their implications understood when making the choice.

Sales Price

The debate over the use of sales price as a measure of property value is quite pervasive throughout the literature. Adkins (1958) states that sales prices are "the best single indicator of economic impact since nearly all effects of highways are eventually relected in land values," while Brigham argues that "sales prices as shown by public records are susceptible to large deviations from the full cash value (of the property)." He found that:

(1) If land value is to be separated from total property value, sales price is useless.

(2) Sales price information as recorded in public records in the form of revenue stamps attached to the recorded grant deed would not be accurate if the purchaser assumes a mortgage loan owed by the seller, because the cash transfer will be less than the actual sales price,

(3) Buyers may overstate the apparent purchase price to help them sell at a higher price later or to "indicate a higher income tax basis" (Although the rationale for the latter reason is unclear, it may indeed be true that the buyer would need to pay less capital gains tax when he sells later).

(4) The seller may take a low-interest rate first or second mortgage in payment for all or part of his equity, in which case the sales price would be well above the full cash value as represented by the revenue stamps.

(5) "Sweetheart sales" (sales to friends or family) would probably result in artificially low or high sales prices.

(6) The last parcel acquired for a large development may sell for far more than the going price in the general area (i.e., a holdout effect).

(7) A "distressed" seller may have to accept considerably less for

his property than he could obtain were he able to delay the sale for a while.

Although these are interesting insights into the accuracy of available sales price data, such insights do not deal with the real issues involved with the determination of property values. The real issue is whether or not the error generated by such deviations in the true sales price is correlated with any of the variables included in the regression equation. If there is no correlation (and most likely there will not be) then such deviations should simply be considered random errors of observations. If there is correlation, another dependent variable should be chosen.

The issue is even clearer in the case of analyzing transit accessibility impacts on property values. Brigham's study was attempting to estimate a model which would accurately predict land values given a set of variables representing characteristics of each parcel; he was thus searching for a regression with a very high R^2, where the independent variables would explain a large proportion of the variation in the data. This means that he would need to check the covariation of the sales price with each and every variable included in the equation. Transit accessibility impact studies, on the other hand, need to look only at the transit accessibility term and any omitted variables which may covary with that term. Thus, in the case of selecting the appropriate dependent variable, one need to look only at whether or not the deviations from true sales price covary with the transit accessibility term.

Tax Assessment Value

Brigham reasons that since land valuation is vulnerable to a great deal of subjectivity on the part of appraisers, it would be desirable to obtain valuations based on a uniform set of premises; one to which all assessors in the area subscribe. This would imply a lower variance, a decrease in the amount of unexplained variation in the regression. Depending on the study area chosen, the premises upon which tax appraisers rely may be satisfactorily uniform. He finds that the Los Angeles County tax assessment system is quite uniform:

> "The California State Board of Equalization takes an active hand in the real estate appraisal practices of the various county assessors. The Division of Assessment Standards of the State Board prepares standard appraisal manuals for use by the local assessors, and this division conducts periodic examinations of the different counties' practices. The Board of Equalization has the responsibility of preparing reports on the various counties' practices including uniformity of appraisals, and the assessors apparently take these reports seriously." (Brigham (1964), p. 83)

He points out, however, that there are three problems with the tax assessment appraisal data:

(1) Properties are assessed in different years, thus introducing complicated and probably uninterpretable lag effects.

(2) Random and systematic errors occur. - He reasons that with a large enough sample size, random errors should average out. However, systematic errors would remain a problem with a strictly cross-sectional analysis. For example, Brigham found that assessors tended to overstate the additional value of a residential property resulting from being located in an attractive neighborhood. This tendency towards an overstatement of the value of properties in "good" neighborhoods would produce a positive correlation between independent variables measuring neighborhood quality and the error in measuring the dependent variable. Consequently, any coefficient for neighborhood quality estimated by

regression on appraisal data would be biased upward.

(3) There may be strong inter-class variation of the percentage of assessed value to appraised market value between different land use types. In other words, the tax assessment rates may be different for different types of land use. Thus, a $50,000 home may be assessed at 50% (or $25,000) and a $50,000 commercial site may be assessed at 80% (or $40,000).

However, as Carroll (1958) points out, "assessed values are determined through a process that covers all properties, whether recently sold or not, and a process which also builds up a total assessment by taking into account the valuation of component parts of the property." (Carroll (1958), p. 47). By covering all properties, assessed values avoid the problem of low turnover rates which is faced by use of sales price data; at least in theory, there is up-to-date data for all properties through time. This may be an important consideration, particularly in the case of time series analysis, when data for each sample unit is required for each year being studied. By taking into account the valuation of the component parts of the property, tax assessment data is able to provide separate land and improvement values, thus facilitating those studies which require only land value or only improvement value data.

Adkins (1958) points out that tax assessments are subject to significant time lag effects, which could have important implications for anaysis of transportation impacts on property values over a fairly short period of time.

Brodsky (1970) used data obtained from a complete file of assessment records for Washington, D.C. 1963 assessments maintained by the Office and Budget and Executive Management in his analysis of the association between residential land values and improvement values in a central city. He found that "generally, assessed values will be a poor source of data for estimating the market of sales value of individual real estate properties." Although he aggregated assessed values by census tracts in order to reduce the random errors, systematic errors such as those experienced by Brigham still persisted.

Appraised Value

Burkhardt (1976) suggests using data supplied by private appraisers, and contends that the data could be generated "much more cheaply and probably with no greater margin of error" than typical statistical techniques. Once a capable appraiser or economist is found, he claims, data about "development pressures at work around a station,...the impact of market, assemblage, and zoning constraints, and...the total value of land value increases within the station area over the next ten to fifteen years" can be generated, purportedly to a very high degree of accuracy.

We have two criticisms of this approach which would discourage its use: (1) the accuracy of a prediction "over the next ten to fifteen years" is subject to human inabilities to predict land development and value changes over such long periods of time, and (2) such an approach is likely to measure what the appraiser believes to be the effect of access to transit on land values rather than the true effect. (If appraisers have an accurate perception of this effect then one could probably avoid any modelling effort by simply asking assessors questions about hypothetical housing units with different access to transit. There is little or no evidence to suggest, however, that appraisers can reliably estimate small differences in property values accurately.)

Brigham does not recommend the appraised value approach for many reasons, all of which stem from the fact that land appraisal is a highly subjective art:

> "In the case of a single-family house, there are a sufficient number of sales of similar properties to permit the appraisers to rely on an established market price: but in the (case of a suburban shopping center or an older apartment building), the appraisers must use subjective judgment in setting values (since) one appraiser may be optimistic about the future, while the other may be relatively pessimistic." (p. 83).

Like Brigham, Carroll (1958) felt that "differences in the purposes of appraisal, and differences among individual appraisers all combine to make the use of appraisal data unsatisfactory for extensively estimating market value on a large scale." (p. 44)

A.II.3 - Appropriate Independent Variables

Other factors (besides the transportation improvement in question) are responsible for changes in property values and must be accounted for in a model of land values; however, the task of selecting the relevant factors is not a trivial one. This is a very fuzzy area, where included variables may be subject to great errors in measurement, proxies, multicollinearities, and where omitted variables may covary with the transit accessibility variable. Such problems with the independent variables may lead to biased and inconsistent estimates of the effect of the transit system on property values.

Some variables which have been used in previous studies to determine property values are:

Wilkinson (1973)

- Average price per square foot
- Distance in miles from CBD
- Number of children under 15 years of age per school
- Number of persons per church in each neighborhood unit
- Quantity of parks, woodland, golf courses

Richardson (1974)

- House type
- Age
- Paid outright
- Distance
- Height
- Direction from central business district (NNE; ENE; ESE; SSE; SSW; WSW; WNW)
- Car ownership
- Zone on periphery
- Class
- No industry
- Rooms per person
- % population under 15
- % population over 65

Hammer (1971)

- Year of sale
- Lot size
- Age of house at time of sale
- Corner dummy
- Irregular lot shape dummy
- Abut-park-directly dummy
- Straight line distance to park
- Distance along public ways to park
- Depth of park/straight line distance to park
- 15 house-type dummies
- 14 year-of-date dummies

Clonts (1970)

- Value of improvements
- Radial mileage to urban periphery
- Distance to urban access highway
- Front feet in residential lots
- Size of parcel
- Sewer and water available
- Poor, fair, good, excellent road conditions
- Date of sale
- Slope of land
- Land subject to flooding

Czamanski (1968)

- Estimated land value per square foot
- Assessed land value per square foot
- Accessibility index
- Lot size in square feet
- Land use category
- Degree of blight
- Age of structures
- Existing zoning
- Zoning change
- Potential change in use
- Ownership

Pendleton (1963)

- Selling price
- Square feet of house
- Square feet of lot
- Construction material dummy
- Basement
- Number of bathrooms
- Extras (garage, carport, recreation room, central air conditioning, etc.)
- Median income
- Age of house
- Number of stories

It is important to point out that as long as we are reasonably sure (intuitively and statistically) that the other independent variables do not covary with the transit accessibility term, then the accuracy of all those other independent variables as predictors of land value is irrelevant to a study such as ours. What we are concerned with is the accuracy of the coefficient of the transit accessibility term. As long as we can determine the effect of improved transportation access on our dependent variable, we have achieved our goal; we are interested in the impact of a transit station on land and improvement values, not in determining the relative effect of all the variables which influence property value.

Given this emphasis on the transit accessibility term, it is critical that in choosing independent variables for the regression equation, the analyst does not omit any variables which might covary with transit access. Say, for instance, that we omit a variable reflecting the quality of a parcel's landscaping from the equation. If indeed people tend to plant more trees and bushes to shield them from the noise and added motor and pedestrian traffic due to the transit station, then the sales price of the property would rise due to the increased accessibility as well as the added landscaping. But since we omitted the landscaping variable, the transit access term would erroneously pick up the entire effect. By omitting just a few covariant variables, it is possible to assign an erroneous coefficient to the transit accessibility term, and thus over- or under-estimate the impact of transit accessibility on property values. So it is essential that every caution be taken to include any influences which may covary with transit accessibility. Although one can never be absolutely certain that all covariant influences are included in the regression, serious efforts to identify such influences should be made.

A.II.4 - Data Sources

Data sources will unfortunately vary from study area to study area, mainly because there is no standard record-keeping procedure for property transactions or tax assessments and because taxing, assessment, and sales procedures vary from geographic area to geographic area.

Contrary to the collection of land use data, which now uses a uniform grid for the classification and mapping of land uses, collecting land value data "...is still a laborious, expensive, and time-consuming process. It is further complicated by the small number of properties changing hands in any one year...The major part of land value changes are the latent or unrealized changes that have not been reflected in the market place or through a commercial sale." (Raup, (1959))

Thus, it is impossible to generalize from prior research and recommend a single, relevant data source for any particular transportation impact study.

However, the following listing of data types and sources used by analysis efforts in the past may prove helpful as a sort of "shopping list" of potential sources:

1) Sales data (market prices of transactions are generally items of public record)

2) Tax assessment value (available from jurisdictional assessor's office)

3) Appraisal value (available by special arrangement from real estate professionals)

4) Tract and block information from U.S. Census on computer tape as well as in bound volumes)

5) Additional residential and work location information is available from the Polk City Directories (St. Paul, Minnesota: Wright Directory Company, 1949, 1959, and 1969).

6) Improvement costs can be determined from:

   - Building permits
   - Local contractor rates
   - Interviews with local residents

In addition, the following sources may provide other useful types of land and improvement value data such as the following (exerpted from <u>BART impact Program: Land Use and Urban Development Project--Phase I Working Papers</u>, November 1975):

1) The Society of Real Estate Appraisers
2) City Planning Departments
3) Building and Housing Departments
4) HUD Survey and Inventory of Unsold New Houses
5) Boards of Equalization
6) U.S. Census of Business and Manufacturers
7) U.S. Geological Survey

A.II.5 - Sampling Procedures

As in the case of choosing an appropriate dependent variable, choosing an appropriate sampling procedure depends on both the object of the research, the available resources and the cost of various options in the area under study.

For the study of unlimited access transportation routes such as major roads, analysts have used successive parallel lines bounding the areas on either side of the routes (Lemly (1959), Adkins (1958)). Data was collected (no specified collection procedures was ever outlined, however) within each band, the property value changes were aggregated for each band, and the results compared relative to each other in a comparative control analytical framework, where land values outside the farthest band were used as control data.

For the study of limited access routes such as freeways, expanding concentric rings drawn around the access points have been used. Data was collected (again, generally with no specified collection procedure) and aggregated within each ring and comparative control analysis was performed in a fashion similar to that of parallel bands.

Clearly, for any study concerned with transit accessibility, distance to the transit access point should be the focus of the sampling technique chosen. In order to increase the statistical precision of a regression analysis, one should strive to maximize the range of the independent variables (in this case, distance to the transit access point), thereby reducing the variance of the regression. One way to ensure this is to stratify the sampling process with respect to distance from the transit access point. This would imply the use of a so-called "ray sampling technique," where a ray is drawn out from the access point and parcels are picked off at determined distances. (Brigham used such a technique in his 1964 study.)

Obviously, the accuracy and usefulness of the analysis would be enhanced by increasing the sample size (within time and budget constraints). Practically speaking, increasing the sample size allows the analyst to deal with problems such as no sales transactions or obsolete tax assessments for observed parcels. In addition, it ensures enough data to permit a stratified analysis of different land use types encountered along the ray.

In order to achieve an operationally optimal sample size, the ray sampling method could be modified in several ways: (1) The ray could be expanded into a narrow rectangle, and any parcels which it encompassed could be used, (2) the number of rays drawn from the access point could vary, (3) the distances from the access point for each ray could vary, and (4) the number of parcels picked off of each ray could vary.

It is difficult to assess the implications of each of these options on the precision with which the effect of transit access is measured. Furthermore, the choice of which stratification rule to adopt will depend on how costly alternative sampling approaches are. Lacking any well-defined theory for choosing among these methods, the best approach is probably to choose the simplest sample design which produces a reasonable amount of variability in the most critical variables (in our case, access to transit).

AII.6 - Appropriate Analytic Methods

There are three types of models which have been described in the literature to analyze the impact of increased transportation accessibility on property values: comparative control, multivariate regression, and factor analysis/orthogonalized regression.

1) Comparative control - studies are based on the assumption that the analyst is able to "hold all other things equal" in order to determine the effect on property value of only the transportation improvement. Thus, study and control areas are observed twice (once before, and once after the transportation improvement), and relative changes in value of the study area in the vicinity of the improvement can be compared with those of the

control areas farther away. As with almost all social experiments, however, the analyst must recognize his/her inability to control the many factors which determine a specific phenomenon. Furthermore, the assumption that one geographical area is in some sense identical to another (except for the existence of the transportation improvement) is unlikely to hold in most cases; this is particularly true in studies evaluating land value shifts, since areas within the same city are part of the same market.

2) Multivariate regression - the most widely used econometric technique, where the analyst utilizes least squares methods to estimate the parameters of multivariate linear predictive models. The motivation for the use of this type of analytical framework lies in the fact that the positive or negative effects of many of the independent variables, and the significance of each of their effects on the dependent variable are unknown, or at best unclear, at the outset of the study. Multivariate regression techniques have the capability of artificially "holding all other influences constant," thus allowing the analyst to estimate the parameters of the model despite the fact that the data does not come from a controlled experiment.

3) Factor Analysis/Orthogonalized Regression - (as it is described by Wilkinson (1973)) is often purported to be useful when it is unclear which independent variables to use, and where there is a significant problem of multicollinearity. (It is important to note that Wilkinson was analyzing the determinants of house prices, and therefore in search of a very high R^2; hence his preoccupation with dealing with a large number of independent variables). We feel that factor analysis is inappropriate in an analysis of the effect of transit accessibility on property values because: (1) the number of potentially relevant variables is reasonable and manageable

within the constraints of the study, and (2) more importantly, factor analysis does not "solve" problems of collinearity; it merely "hides" them by massaging the data into more manageable forms.

For each of the types of models outlined above, there are three main types of data analysis which can be pursued:

Cross-sectional data are observations of a series of properties at the same point in time. The advantages to this approach are that it is generally easier to get data for just one point in time and that one would be able to observe the differing impacts on the value of property as a function of distance from the transit access point. However, there are serious disadvantages to cross-sectional analysis. First, in order for the results of a comparative control analysis to be meaningful, it would be necessary to sample properties which are nearly identical. Such a requirement may simply be impossible to fulfill, given the innumerable variations in property characteristics even within the same land use category in any given area. Second, the selection of an appropriate "control" area which is not impacted by the transit access improvement, and which also is not impacted by any other influence not common to the other parcels in the sample, may be an impossible task; such areas simply may not exist. Third, there is reason to believe that there are significant time lag effects such that the impact of increased accessibility on a property may not be felt until its next sale ot its next tax assessment; thus, making observations at just one point in time ignores these lag effects.

Time series analysis utilizes obervations of the same properties through time. One advantage to this approach is that no "control" areas are needed; the value of each parcel is compared to its value at different times (most likely before and after the transit access improvement). Another advantage is that time lag effects can be observed. Most importantly, in a regression analysis, it is only necessary to include those variables which have changed over time; other descriptors which have remained the same do not need to be included since their effect on the value of the property has been constant through time.

Time sectional data traces a cross section of properties through time. The time sectional approach combines the advantages of both cross-sectional and time series methods. It is a more complete method in that many land use types will be observed as their values change through time: it could be possible to observe varying degrees of the impact of improved transit accessibility on different types of land use. It is also possible to use this approach in a regression analysis and thereby eliminate the property characteristic variables which do not change during the time period being studied. The problem of a control area is also eliminated since each parcel is, in effect, a control for itself over time.

Because our study is based on multivariate regression analysis, we are able to use a sampling approach which offers more flexibility than time series, cross sectional, and time sectional approaches. Multivariate regression allows estimation of the parameters of the model without requiring data which are constrained in any time or cross sectional sense. As a result, it is possible to collect data on different properties at different points in time.

Thus, there are many combinations of methodological elements which have been used to determine the impact of increased transit accessibility on property values. For each method of analysis (comparative control, multivariate regression, and factor analysis), there exist three different approaches (cross-sectional, time series, and time sectional). The following section reviews the research methods advocated by some analysts. While none advocate what we feel is the optimal combination, multivariate regression using unconstrained observations* most of the possible combinations are represented. Some analysts make the effort to discuss the issues involved in determining the appropriate methodology, but often the issues are confused.

Research methods found in the literature

After critically reviewing several papers and studies offering the most relevant analytical contributions in this area of research, Dornbusch advocates the use of multiple regression techniques for two reasons:

(1) It yields the most definitive results (sic),
(2) Its implementation cost is no more than for other less useful statistical analysis methods.

His recommendation suffers from the lack of an operational definition of the criteria for "definitive results." Indeed, a mispecified regression would produce definitive results which would mislead the analyst.

Dornbusch also unequivocally recommends a before/after (i.e., time series) approach instead of a cross-sectional one, based on the assumption that fewer simultaneous variables will need to be included in the model, thus allowing the transportation impact to "emerge more dramatically." Presumably, he means by this that the estimates of the parameters describing access to transit will be more statistically efficient. By using the change in prices as the dependent variable, the

*in a time or cross-sectional

before/after approach requires that only those non-transportation-related influences which have changed since the transit access improvement be included in the analysis; all other influences on land value (size of lot, number of bedrooms, utility availability, etc.) which have not changed do not need to be accounted for.

After analysing three study areas surrounding Texas expressways, Adkins (1959) critically reviews his own comparative control and before/after methodology. He concludes that although such an approach is easy to understand and apparently simple to apply, the approach is not valid since many changes occur to the other factors which influence land values, thus violating the principle of a control. Indeed, Adkins found that an equal matching of control and study areas, the most basic requirement for the comparative control analysis approach, was an impossibility. He was thus forced to compromise the methodology and proceed with "reasonably comparable" study and control areas.

Cribbins, Hill, and Seagraves (1964) concur with Dornbusch and use the before/after approach to multiple regression analysis for three sections of interstate routes in North Carolina.

Lemly (1959) uses both time series and cross-sectional data with the comparative control method to determine changes in land use and value along segments of expressway in Atlanta, Georgia. He was thus able to compare study and "control" areas both before and after construction as well as compare study areas of one type with study areas of another type to determine the degree of influence and the timing of influence of expressway construction. He did not, however, mention how appropriate his choices for control areas were, and the extent to which the factors which influence those areas (besides the transportation improvement) remained constant over time.

In his study of the interrelations between transportation facilities and the location of economic activities in the Los Angeles areas, Brigham (1964) used the cross-sectional, rather than a before/after approach "primarily because reliable time series information is difficult to obtain." This may be true, particularly in the case of his investigation; however, considering the limitations of the use of cross-sectional data as discussed earlier, it seems questionable whether that difficulty is more serious than the unreliability of cross-sectional analysis of land values in an urban area.

In their study on the measurement of the determinants of relative house prices, Wilkinson and Archer (1972) test the appropriateness of both the multiple regression and the factor analysis/orthogonalized regression methods. They find that the regression technique could not accomodate the many different dimensions to the housing environment; i.e., that no single regression equation adequately described the land pricing phenomenon since each one reflected influences beyond those that it was intended to measure. Meanwhile, they claimed that their experience with factor analysis of the same data set proved much more favorable due to the fact that such an approach "tests the usefulness of individual variables." Again, we feel that factor analysis merely hides collinearity problems.

In response to Wilkinson and Archer, Richardson, Vipond and Furbey (1974) advocate the use of multiple regression analysis, not factor analysis. They reason that factor analysis is more appropriate if there is little justification for choosing a few key variables out of a great pool of variables. However, in their study of the determinants of urban house prices, they found that it was possible to theoretically justify the inclusion of each variable used.

Summary of Methods

Very few of the past works are able to offer truly definitive conclusions on the relative importance of transportation improvements on property values. Indeed, the state-of-the-art of measuring impacts of various factors on property values, while progressing, is still at a very low level of development. Furthermore, there is no consensus in the literature as to the "correct" methodology, although the trend seems to be heading towards the multiple regression techniques. Indeed, it may be generalized that studies conducted in the 1950's and early 1960's tended to be of the comparative control genre, while those conducted in the late 1960's and early 1970's are of the multiple regression genre.

A.II.7 - Conclusions About Elements of Analysis

A general conclusion which may be drawn from the literature is that the choices of analytic method, dependent variable, independent variables, sampling procedures, and data sources are all highly dependent on the study site selected. In other words, analysts have found no specific methodology which is appropriate for application to the general problem; indeed it seems that for each site to be studied, a new methodology and data collection procedure must be developed, tailor-made to the characteristics and demands of the site.

Thus, it seems most appropriate to offer the following review as a guide to developing a methodological procedure given particular study site characteristics:

(1) Scope and types of study areas

Government jurisdictional boundaries are often indicative of the boundaries of tax assessment methods, zoning regulations, etc. The type and service characteristics of the particular transportation improvement in question may often dictate the scope of the study area.

(2) Choice of dependent variable

Before choosing the appropriate dependent variable, it is necessary to decide on the question of defining "property value" as including both land and improvements or simply land. If the method of analysis used is comparative control, it would be advisable to use only land value, since it can be expected that over any period of time, there will be changes in the quality of improvements (positive or negative), and since the comparative control method is incapable of dealing with dynamic changes in control and study areas (Note then that if data on land value only is not available, it would be virtually impossible to conduct a credible comparative control analysis.). As is previously noted, several studies (e.g., Brigham (1964)) were concerned about this question. The debate seems pointless, however, if the method of analysis used is multivariate regression since multivariate regression is capable of "isolating" and "controlling" the variables affecting improvements as long as one has data on improvements.

Assuming that most contemporary studies will be more concerned with multivariate analysis and land improvement values, one could use:

(i) Sales price - if information about the total sales price (not just cash price) is available; if there is a high property turnover rate in the area which will generate a sufficient data base, or

(ii) Tax assessment value - if all properties are assessed in coincident years, if the procedures used by local assessors minimize systematic inter- and intra-class biases of individual assessors; if the study methodology can deal with the significant time lag effect of assessment.

(iii) Appraised value - if the study area is relatively small (manageable) if capable appraisers or economists who know the study area very well (in terms of zoning, developmental pressures, general market conditions, etc.) can be obtained.

The resolution to this issue depends on what one wants to measure. Thus, if the purpose of the analyst's study is to determine how improvements in transit access affects tax revenues, the analyst should use tax assessment values; if the analysis concerns how appraisers feel about the effects of improvements in transit access, the analyst should use appraised values; finally, if the purpose of the analysis is to measure how the market values of properties respond to improvements in transit access, then the analyst should use sales price.

(3) Choice of independent variables

This is a crucial step in any analysis, for the inclusion of too many insignificant independent variables can overcomplicate the regression to the point of rendering it useless, while the exclusion of significant ones can result in mis-specification of the model. Each study effort should carefully consider all independent influences which have been deemed significant in past studies (there is a condensed list extracted from several past studies included earlier in this text) and come to a decision on their inclusion into the relevant regression equation.

Although the inclusion or omission of specific variables is a serious and critical problem, it must also be emphasized that it is one thing to attempt to predict the true value of property as a function of several independent variables and yet quite another thing to isolate the true effect of transit accessibility on property value. In the extreme cases, assume all other factors determining property value are entirely uncorrelated with transit accessibility. Then a regression analysis would yield unbiased and statistically efficient estimates (i.e., $E(\hat{\beta}) = \beta$ and $var(\hat{\beta})$ is minimized) of the coefficient of transit access even if all the other variables are omitted from the model. (Obviously, the estimated equation would yield very imprecise estimates for the dependent variable.) The analyst should thus be aware that a high or low R^2 for the regression equation is virtually meaningless when studying the effect of transit access on property values.

Two techniques have been offered as means of selecting the appropriate set of independent variables:

(i) Factor analysis is discussed in cases in which there is little justification for selecting a few key variables from a large number of variables. As discussed previously, the number of independent variables involved in the determination of property values is quite large but manageable, and such mechanistic statistical methods may introduce variables which impose a set of linear constraints on the model which may be inappropriate.

(ii) Stepwise linear multiple regression has been proposed when there is a manageable number of variables which are thought to be significant. This type of regression procedure claims to ensure that any variable which provides an insignificant contribution is removed from the model. There are, however, very serious shortcomings to the process of stepwise linear regression, the most obvious of which is that the variables which are chosen through this process may appear in the final model simply because they were considered first. Suppose two independent variables, X_1 and X_2, covary with each other. Any common influence of these two regressors on the dependent variable would be attributed to the first one chosen, thus "robbing" X_2 of its true effect (see Wonnacott and Wonnacott, pp. 309 - 312 for a more detailed discussion of the limitations of stepwise linear procedures.) In the case of estimating the impact of transit access on property values, if the transit access term were X_1, its impact on property values would be overestimated; if the transit access term were X_2, its impact on property values would be underestimated (Dornbusch (1975), p. 6 briefly discusses this point.)

(4) Data sources

Those data sources which are chosen must fulfill at least the following requirements:

(1) They must be complete (in terms of having information for each sample) over the time period under study,

(2) They must be available to the study group.

(3) They must be uniform over the geographic area under study (e.g., tax assessment procedures which are uniform over two or more jurisdictions being studied).

(5) Sampling procedures

The issue of which sampling procedure is appropriate is one which concerns parameter efficiency traded off with cost functions and constraints unique to each study. The tradeoff is a difficult one. In economic terms, the size of the sample is determined at the point where the marginal "product" of the last parcel sampled to parameter efficiency just equals its marginal cost of collection. Unfortunately, while the marginal cost of collection can be estimated quite easily, the marginal cost of the higher variance due to not collecting is rather difficult to ascertain.

(6) Appropriate analytical methods

Three main methods of analysis and three distinct approaches to data sampling form the core of this key decision in the modelling process. Comparative control, multivariate regression, and factor analysis/orthogonalized regression are the three analytical frameworks which have been used by analysts in the past. The types of data which have been used within these frameworks are cross sectional, time series, and time sectional. Almost any combination of framework and data is compatible. However, the optimal combination for the study of the impact of transit access on property values, in our view, is the use of pooled time series and cross-sectional data within the framework of a multivariate regression analysis. Such a methodological decision is based on the unavailability and unreliability of controls (in any time or space dimension) and on the availability and reliability of data unconstrained by time series or cross-sectional requirements; and most importantly, on a clear definition of the ultimate goal of the analysis.

A.III. HYPOTHESES GENERATED FROM THE LITERATURE

A survey of the literature has also produced a set of tentative hypotheses about the determinants of land values. In some cases, these hypotheses are simply the result of speculations or casual observation by the author; in other examples, the hypotheses were actually tested in an empirical study. These hypotheses will provide the basis for specifying the regression model and comparing the estimated coefficients with previous results. Listed below are the major hypotheses generated in the course of the review:

(1) Decreases in value elsewhere occur which offset rises in value associated with the transportation improvement, i.e., the increase in total aggregate land value is small compared to the redistribution of value. (Beesley and Foster (1973), Allen and Mudge (1974), Burkhardt (1976), Gannon (1975), Goldberg (1970), Mohring (1961)).

(2) The existence and extent of property taxes themselves are depressing factors on the property value (Boyce (1972)).

(3) Travel time to the CBD and other significant activity centers is capitalized into property values; thus, time savings will consequently result in value increased (Allen and Mudge (1974), Boyce (1972), Dornbusch (1975), Waldo (1974)).

(4) Transit improvements act primarily to redistribute development within a region, not create it (Burkhardt (1976), Czamanski (1966), Goldberg (1970)). The theory is that the resources of a region

can support just a certain level of development, and that perhaps all that a transportation system can do in terms of development is affect the location of that development.

(5) Properties too close to the station or surface part of a line suffer value decreases due to nuisance level increases (Burkhardt (1976), Dornbusch (1975)).

(6) Areas which are already developed do not generally show a marked increase in land value after a new line is opened (Spengler (1930)).

(7) Areas which are already supplied with transportation lines will only experience moderate changes in land values (Spengler (1930)).

(8) Property value increases are subject to large time lags (Adkins (1958), Carroll (1958), Langfeld (1971), Rogers (1963)).

(9) General increases in land value are determined by other forces in the area which may be interrelated with the facilities (Ashley and Bernard (1965), Beesley and Foster (1973), Corsi (1974), Cribbins, et. al. (1964), Gauthier (1970), Lee (1973a), Spengler (1930)). For example, commercial and shopping development may follow the construction of a transit facility, and proximity to those developments, as well as to the transit access point, may influence the value of property.

(10) Real estate values should increase in direct relation to

diminishing risk to developers and landlords through elimination of uncertainties (Davis (1970), Lee (1973a)).

(11) Real estate values should increase with diminishing time before the transportation system is operable (Davis (1970)).

(12) General accessibility is highly important for local commercial activities (Downing (1973)).

A.IV. SUMMARY OF CONCLUSIONS

This survey of the literature has produced a number of conclusions pertaining to the development of research methodology with respect to the analysis of transit accessibility impacts on property values. In summary, the most relevant of these conclusions are as follows:

(1) There seems to be no single, appropriate methodology for choosing the appropriate independent variables. Although several statistical techniques such as stepwise linear procedures and factor analysis have been proposed in the literature, use of such methods creates serious problems of misspecification and "hiding" of potentially valuable information.

(2) Very few studies have pursued research in this area in a very analytically rigorous fashion. Those which were particularly sloppy in their methodological decision-making (sample size, appropriate dependent and independent variables, data sources, overall analytical framework) tended to come out with "definitive" conclusions: i.e. hard numbers to justify their hypotheses. However, when considering their methodologies, the accuracy of their models and reliability of their predictions seemed highly questionable. On the other hand, those studies which took pains to make the careful methodological decisions tended to qualify their findings and seemed unable to commit themselves to any "definitive" conclusions on the basis of their empirical findings. This latter case illustrates the difficulty of this type of research and the elementary level of the state-of-the-art.

(3) The nuisance effects associated with transportation facilities have rarely been considered by analysts in the past. But with multivariate regression analysis, it is possible to incorporate variables which reflect the level of negative effects of in the model.

(4) When using any type of regression technique, the key issue is not to omit any variables which may covary with the transit access term; by omitting such variables, one leaves the transit access variable to "pick up" all the effects, thus over-or under-estimating the true effect of transit accessibility.

(5) When using regression for the analysis of impact of transit access on property values, it is of secondary importance to achieve a high R^2, so long as the transit access term and any omitted terms are not highly correlated.

(6) Any methodology which relies heavily upon controls in the sample data (in any time or space dimension) is likely to be subject to problems since the real world has a tendency to violate any restrictions which social scientists put on their observations. On the other hand, methodologies which utilize statistical controls <u>within the model</u> (i.e., multivariate regression techniques), enable the analyst to isolate and estimate the true effect of the independent variables on the dependent variable.

A.V. LIST OF AREAS STUDIED

AREAS	DATE	AUTHOR	TITLE
ATLANTA, GEORGIA	1959	Lemly, J.H.	"Changes in Land Use and Value Along Atlanta's Expressways."
	1976	Burkhardt, R.	Joint Development Study.
AUSTIN, TEXAS	1960	Wootan, C.V. & C.R. Haning	Changes in Land Value, Land Use, and Business Activity Along a Section of the Interstate Hwy System in Austin, Texas.
BALTIMORE, MARYLAND	1976	Burkhardt, R.	Joint Development Study.
BOSTON, MASSACHUSETTS	1973	National League of Cities.	Transit Station Joint Development.
BUFFALO, NEW YORK	1973	National League of Cities	Transit Station Joint Development.
CHICAGO, ILLINOIS	1968	Golden, J.S.	"Land Values in Chicago: Before & After Expressway Construction."
	1973	National League of Cities	Transit Station Joint Development.
DALLAS, TEXAS	1958	Adkins, William G.	"Effects of the Dallas Central Expresswa on Land Values and Land Use."
DENVER, COLORADO	1976	Denver, Colorado Regional Transportation District	"Value Capture Opportunities."
GLENDALE, CALIFORNIA	1971	Miller, S.G.	"Effects of Proposed Highway Improvements on Property Values."
GREENSBORO, N.C.	1963	Rogers, Andrei	Time Lag of Factors Influencing Land Development.
HOUSTON, TEXAS	1959a	Adkins, William G.	"Land Value Impacts of Expressways in Dallas, Houston, & San Antonio."
I-94, MICHIGAN	1965	Ashley, R.N.	"Interchange Development Along 180 Miles of I-94."
LEXINGTON, KENTUCKY	1975	University of Kentucky	The Effect of the Louisville-Watterson Expressway on Land Use and Land Values.
LINDENWOLD, NEW JERSEY	1972	Boyce, David	"Impact of Rapid Transit on Suburban Residential Property Values and Land Development."
	1974	Allen, Bruce	"The Impact of Rapid Transit on Urban Development--Case of the Philadelphia Lindenwold High Speed Line."
	1975	Gannon, Colin A. & Michael Dear	"Rapid Transit and Office Development"

A.V. LIST OF AREAS STUDIED (Cont.)

AREAS	DATE	AUTHOR	TITLE
LONDON, ENGLAND	1973	Beesley, M.E.	"Estimating the Social Benefit of Constructing an Underground Railway in London."
LONG BEACH, CALIFORNIA	1961	Huhtanen, Robert J.	A Study of the Effects of Freeways on CBD's.
LOS ANGELES, CALIFORNIA	1964	Brigham, E.F.	A Model of Residential Land Values.
	1974	Waldo, Robert D.	Urban Land: Values and Accessibility.
MASSACHUSETTS	1973	National League of Cities	Transit Station Joint Development.
MICHIGAN	1965	Ashley, R.N. & W.F. Bernard	"Interchange Development Along 180 Miles of I-94."
MILWAUKEE, WISCONSIN	1973	Downing, P.	"Factors Affecting Commercial Land Values: An Empirical Study of Milwaukee, Wisconsin."
	1973	Soot, Siim	"Transit Costs and Urban Land Rent Theory The Milwaukee Example: 1949-1969."
MINNESOTA	1958	Carroll, Donald D.	The Economic Impact of Highway Development Upon Land Use and Value.
NEW YORK: (Syracuse Utica-Rome Albany-Schenectady-Troy Buffalo Binghamton)	1971	Barden, R. & J.H. Thomson	The Urban Frontier.
NEW YORK CITY	1917	Cushman, R.E.	"Excess Condemnation."
	1930	Spengler, E.H.	Land Values in New York in Relation to Transit Facilities.
NORTH CAROLINA INTERSTATE ROUTES	1964	Cribbins, P. D.	"Economic Impact of Selected Sections of Interstate Routes on Land Value and Use."
OAKLAND, CALIFORNIA	1961	Huhtanen, Robert J.	A Study of the Effects of Freeways on CBD's
	1976	Burkhardt, R.	Joint Development Study.
OHIO TURNPIKE	1974	Corsi	"A Multivariate Analysis of Land Use Change: Ohio Turnpike Interchanges."

A.V. LIST OF AREAS STUDIED (Cont.)

AREAS	DATE	AUTHOR	TITLE
RICHMOND, VIRGINIA	1961	Huhtanen, Robert J.	A Study of the Effects of Freeways on CBD's.
ST. LOUIS, MISSOURI	1971	Miller, S.G.	"Effects of Proposed Highway Improvement on Property Values."
SAN ANTONIO, TEXAS	1959b	Adkins, William G.	"Economic Effects of Expressways Without Paralleling Service Roads in the City of San Antonio."
SAN FRANCISCO BAY AREA	1970	Anderson, Arnold C.	"The Effect of Rapid Transit on Property Values."
	1970	Davis, Frederick W.	"Proximity to a Rapid Transit Station as a Factor in Residential Property Values"
	1972	Lee, Douglas B., Jr.	"Analysis of BART Impacts on Bay Area Land Use."
	1975	Dornbusch, D. M.	"Transit Access-Induced Changes in Property Values and Rents"
	1976	Boyce, Allen	BART Working Papers.
SEATTLE, WASHINGTON	1961	Mohring, H.	"Land Values and the Measurement of Highway Benefits."
TORONTO, ONTARIO	1968	Heenan, G.W.	"The Economic Effect of Rapid Transit on Real Estate Development."
	1971	Langfeld, S.C.	"The Balanced and Orderly Development of the Site in Close Proximity to a Metro Station."
	1973	Deewes, D.N.	The Impact of Urban Transportation Investment on Land Value.
	1976	Davies, Gordon W.	"The Effect of a Subway on the Spatial Distribution of Population."
VANCOUVER, B.C.	1970	Goldberg, Michael A.	"Transportation, Urban Land Values, and Rents: A Synthesis"
WASHINGTON, D.C.	1959	Burton, Robert C.	"Socio-Economic Change in the Vicinity of Capital Beltway in Virginia."
	1963	Pendleton, W. C.	"Relation of Highway Accessibility to Urban Real Estate Values"
	1970	Brodsky, H.	"Residential Land and Improvement Values in a Central City."
	1970	Davis, Frederick	"Proximity to a Rapid T Station as a Fact[illegible] in Residential Property Values."
	1971	Langfeld, S.C.	(see under Toronto, Ontario)
	1976	Burkhardt, R.	Joint Development Study.
YORK COUNTY, PENNSYLVANIA	1965	Eyerly, R. W.	Property Formation, Real Estate Value & Land Use in Four York County Interchange Communities.

A.VI. TABLE OF LITERATURE ABSTRACTS

The following table presents all the works abstracted in the annotated bibliography. The categories are described as follows:

AUTHOR: first name cited in bibliography, in the case of more than one author.

DATE: publication date. Followed by a, b, c in the case of more than one publication by the same author in the same year (the a, b, c corresponds to the coding of the following bibliography).

OBJECTIVE: the stated objective of the publication may not necessarily correspond with the objective of the analysis of transportation improvement impacts on land value, although parts of the publication may be relevant.

TYPE OF FACILITY: the type of transportation facility studied in the literature varied from expressway interchanges to free-access highways, to BART stations to the Lindenwold fixed heavy rail corridor, and are so noted in the table.

TYPE OF ANALYSES: the methods of analysis used by the authors included comparative control, time series and cross-sectional multivariate regression, ANOVA, and other methods including field surveys and historical review.

DATA SOURCES: the type of data used in each analysis may be inferred from the source noted.

DEGREE OF DETAIL: a general indicator (Low, low-med, medium, med-high, high) of the detail into which each analysis goes.

DEGREE OF RELEVANCE: a general indicator (same as above) of the relevance of each publication, or a part of each publication, to the analysis. of transportation improvement impacts on property values.

RELEVANT CONCLUSIONS: conclusions which are relevant to the analysis of transportation improvement impacts on property values; not necessarily the major conclusions of the publications.

COMMENTS: brief mention of the significance of the publication with respect to our project, and/or the quality of the publication with respect to the methodological aspects of this type of research.

Those abstracts with asterisks are summarized from the following literature reviews:

(*) Burkhardt, Ross, Summary of Research: Joint Development Study, Administration & Managerial Research Association of New York City, Mayor's Midtown Action Office, 220 W. 42nd St., 23rd Fl., New York City, NY 10036, 1976.

(**) U.S. EPA, Secondary Impacts of Transportation and Wastewater Investments: Review & Bibliography, Office of Research & Development, Socio-Economic Environmental Studies, January, 1975.

A.VII - Annotated Bibliography

AUTHOR	DATE	OBJECTIVE	TYPE OF FACILITY	TYPE OF ANALYSIS	DATA SOURCES	DEGREE OF DETAIL	DEGREE OF RELEVANCE	RELEVANT CONCLUSIONS	COMMENTS
Adkins	1959	T improvement impact on land values	Expressways	Before/after; comparative control	Interviews, field observations, maps, photos	High	Medium	Land closest to T increased in value 300-600%, land farther had smaller rises.	Section on methodology is excellent
Allen & Mudge	1974	T improvement impact on land values	Fixed heavy rail		Real estate transactions	Med-high	Med-high	Increases within T area balanced by decreases outside the area.	Important conclusions
Anderson	1970	Hard-sell BART	Fixed heavy rail	none	Anderson's own estimates	Nil	Nil	None	No relation between title and content article
Ashley	1965	Interchange development and land value increase	Highway interchanges	Before/after		Medium	Low-Med	Issue is not whether increases exist, but how great they are.	Methodology not particularly relevant, but conclusions are important to note
Averous & Lee	1973	Define balance between T and productive activity	None	Static equilibrium model		High	Low	T is both provider of access and consumer of land	Content of article not relevant to project
Barden	1971	Determine major factors in urban growth	None	(Unspecified) model	Real estate values	Medium			Methodology and model used interesting
Beesley & Foster	1973	Estimate social benefits of T	Underground railway			High	Low-med	Increases in land values are often false measure of benefits	Interesting conclusion about treatment of land values & measure of benefits
Boyce & Allen	1976	Discuss analysis of T impacts	Fixed heavy rail	(1) Exploratory data analysis (2) Statistical hypothesis testing (3) Urban model building	None	Medium	Low	(1) and (2) are more suited to impact analysis; (2) is "too coarse to be useful"	Puts model building into good perspective
Brigham	1964	T and location of activities	Road network	Multiple regression; cross-sectional	County assessor's data	High	High	Model confirms + relation, but not accurate enough for individ. prediction.	Study methodology well-thought out and executed.
Brodsky	1970	CBD improvements impacts on residential land values	CBD real estate	Statistical analysis; aggregate data	Assessment data	High	High		Discussion of statistical methodology is excellent
Burkhardt	1976	Literature search & T impacts on land values	Various	Literature search & real estate consultants	Literature & real estate consultants	High	Very high	+ land value impact depends on additional factors	Excellent, up-to-date work
Burton	1959	Econ. and social changes due to T improvement	Automobile beltway	Time series; comparative control	Sales data	High	Medium	Upward trend of land values reflected by sales data	Good analysis; significant conclusions; but note, this is a highway agency report, so conclusions may be subject to bias
Carroll, et.al.	1958	T impacts on land value and use	Highway	Multiple regression	Ratio of assessed value of buildings/land	High	Med-high	+ relation inferred for land value from higher-intensity land use	Study area selection methodology is relevant
Coral	1974	Land use changes	Turnpike interchanges	Multivariate analysis	Aerial photos	Med-high	Low-med	Growth around interchanges is highly variable	Good example of multivariate regression technique; however, deals only with LAND USE CHANGES, not VALUE CHANGES due to T

A.VII.-Annotated Bibliography (Cont.)

AUTHOR	DATE	OBJECTIVE	TYPE OF FACILITY	TYPE OF ANALYSIS	DATA SOURCES	DEGREE OF DETAIL	DEGREE OF RELEVANCE	RELEVANT CONCLUSIONS	COMMENTS
Cribbins, et.al.	1964	T impacts on land values and development	Interstate routes	Before/after; multiple regression; computer analysis	Sales prices	High	High	Significant value increases not attributable to T improvement	Excellent discussions of multiple regression technique performed and method of land value data collection
Czamanski	1966	Development policy impacts on land values	None	Multiple regression ANOVA	(Capitalized value)-(depreciated replacement cost)	High	Medium	Accessibility to CBD is prime determinant of value	Main conclusion speaks well for notion that land values increase with T improvement
Davies		T impact on spatial dist^n of population	Subway	Regression analysis	Gross density data from U.S. Census	High	Low-med.	+ change in pop. greater near stations; time lag between constr. and devel.	Value changes probably faster than use changes. Discusses only land use, not value
Davis	1970	T impact on land value	Fixed heavy rail	"Map analysis"	Sales prices and maps	Low-medium	Med-high	+ price trends due to T improvement	Relevant study area selection method
Dornbusch	1975	Methodology for analysis of T impacts on land values	Heavy fixed rail	Before/after analysis; Multiple regression	Property sales tax	High	High	Various methodological recommendations	Excellent reference for methodology for study of T impacts on land values
Downing	1973	T impact on land values	Bus routes	Statistical analysis		Med-high	Low-med	Land values highly dependent on accessibility	Conclusion confirms notion that land values are dependent upon level of accessibility
Eyerly	1965	T impact on land values	Highway interchanges		Tax record data	Medium	Low-med	Interchange value increases LESS than surrounding areas	Note the use of tax record data. Important conclusions.
Gannon	1975	T impact on office development; generate hypotheses for further research	Fixed heavy rail	Field survey	Users and suppliers of office space	Low	Low	T has significant impact on spatial structure of commercial office market	No rigorous analysis; but hypotheses generated summarize current issues concerning urban analysts
Gauthier	1970	T in developing Countries	Railroads	Historical review	Literature	Med-high	Low	+ changes at expense of other areas	Cites historical events which indicate how T led development
Goldberg	1970	T, land values, and rents	Fixed heavy rail			Medium	Med-high	T accessibility will lead to decrease in aggregate land values	Excellent synthesis of past work on aggregate effects of T improvements
Goldberg	1972	T impact on land values	Freeway		Deflated price index of randomly selected properties	High	Low	Land 0.1 mi. from freeway increases in value; others constant or decrease	Important conclusions. Note separation of land from land improvements
Golden	1968	T impact on land values	Expressway	Statistical analysis Comparative control		High	Med-high	Value increases greater than control areas; land near freeway not always higher	Important conclusions. Note use of control areas
Hamburg	1970	T impact on land development	None	Prototype T and land use development model	None	High	Low	T shapes urban development	+T impact on land values may be inferred from +T impact on land density
Hammer	1971	Large Park impact on real estate values	Large urban park	Multiple regression	Land value only	High	Med-high	Park coefficient significantly large	Multivariate regression discussion is elementary. Faulty rationale for using land value, not total prop. value
Heenan	1968	T impact on land values	Fixed heavy rail	Simple reporting of present land values	Assessed value	Low	Medium	Values increased more rapidly than elsewhere in city	Lack of satisfactory control area is a significant drawback to this analysis

A.VII - Annotated Bibliography (Cont.)

AUTHOR	DATE	OBJECTIVE	TYPE OF FACILITY	TYPE OF ANALYSIS	DATA SOURCES	DEGREE OF DETAIL	DEGREE OF RELEVANCE	RELEVANT CONCLUSIONS	COMMENTS
Huhtanen	1961	I impact on land use, value, access, travel patterns	Expressways	Land use mapping techniques	Land use maps	Medium	Low-med		Note the use of land use maps
Talbor	1975	I impact on land value	Highway	Step-wise multiple regression			Low-med		Type of highway, type of land use, and location of parcel are significant variables
Kentucky, University of	1975	I impact on land value	6-mile free-access bypass	Comparative control	Land values, case studies, surveys of comm. and ind. establishments	Med-high	Low-med	Significant + changes in land values	Important conclusions; note the use of comparative control method
Langfeld	1971	I station impact on land values	Heavy fixed rail	Simple observation	Assessed values	Low-med	Low-med	+ changes in values; lag time between market value and assessed value noted	Certainly not a rigorous analysis; time lag effect is important to note
Lee	1972	Research design for I impact on land use	Heavy fixed rail	Descriptive, behavioral, normative	None	Med-high	Med-high	Strong theoretical model needed, but difficult to develop	Good summary of state-of-the-art of I impact on land value analysis
Lee	1973a	I impact on land values	Heavy fixed rail	Longitudinal and cross-sectional analysis	Property value data	High	Med-high	+ changes also dependent on other factors besides I	Note use of both time series and cross-sectional analysis. Important conclusions
Lee	1973b	Normative framework for I policy w.r.t. land use	None	Describes 4 models	None	Medium	Medium	I leads to higher intensity development and higher property values	Clear, simple models of basic relationship between land use and I. Note conclusion.
Lemly	1959	I impacts on land use ad value	Expressway	Comparative control	Real property sales data	Medium	Med-high	+ changes in most land near expressway	Study does not take redistribution of value increases over entire area into account
Miller	1971	I impacts on land value	Highway	Case studies	Land value/acre	Low	Low	+ change at interchange from $400/acre to $14,500/acre	Not rigorous; merely reports land value changes over time for specific sites
Mohring	1961	Use of land values in analyzing I impacts	Highway	Extensive theoretical and empirical analysis.	Land value data	High	High	Increases in values of land near highway balanced by relative decreases elsewhere	Develops good argument that I does not necessarily increase aggregate land values in a region
Pendleton	1963	I impacts on land values	Highway	Regression analysis; aerial photos	Sales prices	High	Med-high	Use sales data to estimate accessibility values	Good justification for using sales data to estimate accessibility values
Richardson	1974	Determinants of urban house prices	None	Multiple regression; aggregate data	None	High	Med-high	Use hybrid theories to derive independent variables	Unique research recommendations.
Rogers	1963	Time lag of factors influencing land development	None	Multiple regression; time series		Med-high	Low-med	Work and school accessibility significant variables	Note the significance of distance to elementary school variable
Seyfried	1963	Centrality of urban land values	None	Regression analysis	Assessment records	High	High	Market forces allocate supply of sites and values	Main thrust of piece not relevant, but data collection techniques useful
Spengler	1930	I impact on land values	Subway		Assessed value data	High	High	I SHIFTS value; does not CREATE; land values determined largely by other factors	Perhaps the most widely-quoted work in existing literature. Significant conclusions.

A.VII - Annotated Bibliography (Cont.)

AUTHOR	DATE	OBJECTIVE	TYPE OF FACILITY	TYPE OF ANALYSIS	COMMENTS	DATA SOURCES	DEGREE OF DETAIL	DEGREE OF RELEVANCE	RELEVANT CONCLUSION
Waldo	1974	Measure the time value of commuting as it affects land rents	Road network	Random sampling procedures	Accessibility-land use relationship discussed, using land values. Does not address the T impact on land values.	Land rents	High	Low-med	Verifies accessibility-land use relationship
Wooten & Haning	1960	T impacts on land values and use and business activity	Interstate highway	Comparative control	Not relevant since it considers undeveloped areas while we are concerned with value changes in already-developed areas.	Land value	Medium	Low	Land values related to land use and distance from CBD's

POLICY REFERENCES

AUTHOR	DATE	OBJECTIVE	TYPE OF FACILITY	TYPE OF ANALYSIS	COMMENTS	DATA SOURCES	DEGREE OF DETAIL	DEGREE OF RELEVANCE	RELEVANT CONCLUSION
Cushman	1917	Foreign experience with value capture	Various	Description, Historical survey	Although an old work, insights are quite relevant to the contemporary arenas.	Historical references	Medium	Med-high	Good administration, favorable market conditions are prerequisites for making a profit from supplementary condemnation
Denver, Colorado	1976	Outline specific value-capture policies	Rapid transit	Description	Excellent summary of various value capture techniques.	None	Low-med	High	Value capture policies can be successfully applied
Grimes	1975	Review foreign experience with value capture policies	Various	Historical survey	Most of conclusions are intuitively appealing.	Land taxes, incremental taxes	Med-high	High	Foreign experience not overwhelmingly successful
Lowenberg	1974	Considers means by which increases + decreases in land values may be captured or mitigated by the public	Various	Analytical description of methods	Excellent reference for types of VC policies and their effectiveness.	Methodology references	High	High	Windfalls for wipeouts?
Sharpe	1974	Legal, community enhancement, + financial aspects of VC are explored	Fixed guideway	Documented case studies + related research methods	Very detailed, good overall summary of several aspects of VC policies.	Historical references	High	High	VC is an innovative method for encouraging T improvements

BIBLIOGRAPHIC REFERENCES

AUTHOR	DATE	TOPIC	DEGREE OF DETAIL	DEGREE OF RELEVANCE	RELEVANT CONLUSIONS	COMMENTS
Burkhardt	1976	Land value windfalls in station areas.	High	High	Appraisal and market analysis may generate the needed information more cheaply than statistical techniques. Land value impact of a T station is dependent on other factors.	Very thorough review; excellent reference.
Dornbusch	1975	Transit access--induced changes in property values and rents.	Low-Med	High	Develops a methodology for determining property value impacts of T accessibility.	Although he only reviews a few sources, his criticisms of these sources are quite relevant and useful for developing a study methodology.
Lynch	1975	T and Spatial Form: Land use and development impacts.	Low	Low	* * *	There are few references in this bibiliography which relate to property value impacts.
Onibokum	1975	Socio-economic impact of highways and commuter rail systems.	Med	Med-High	The goals of the study effort must dictate the appropriate methodology, study areas, data sources, etc.	A very relevant bibliography, although the annotations were not in themselves very thorough. No conclusions were offered.
U.S. EPA	1975	Secondary Environmental impacts on various forms of public investments.	Med-High	Med	The appreciation of urban land depends on its location with respect to the highway, and value of single-family dwellings does not, on average, increase with greater highway accessibility.	Excellent reference. Its conclusions concur with those of our literature search. It covers a broader topic range, and hence the relevant section on impacts of transit access is rather small.

The effect of the Washington
Metro on urban property