The Demand for Residential Telephone Service
The Demand for Residential Telephone Service
The Demand for Residential Telephone Service

Gary P. Mahan

1979
MSU
Public Utilities Papers

Division of Research
Graduate School of Business Administration
Michigan State University
East Lansing
Contents

List of Tables
Acknowledgments

Chapter 1
Introduction

Objectives 1
Usefulness of Results 2
Improved Planning and Marketing by Telephone Companies 2
Improved Regulation of Telephone Companies 4
Weaknesses of Prior Research 6
Scope and Limits of Study 7
Description of Study 8
Chapter 2
Review of Previous Studies

Stern 13
Dobell et al. 13
Gary, et al. 17
Long Distance Interstate Message Toll Service Demand Model 19
Irish 22
Perl 26
Survey Research Center Study 29
Difficulties Encountered in Analyzing Aggregate Data 31

Chapter 3
Demand for Residential Long-Distance Telephone Service

Theoretical Model 35
Unit of Analysis 36
Dependent Variable 36
Income 38
Measurement Problems 39
Theoretical Problems 40
Functional Form 42
Prices 45
Traditional Noneconomic Variables 45
Household Size and Composition 46
Degree of Urbanization 49
Social Class 50
Race 51
Special Noneconomic Variables 52
Geographical Dispersion of Family 52

Chapter 4
Demand for Residential Extension Stations

Theoretical Model 88
Economic Variables 88
Traditional Noneconomic Variables 89
Special Noneconomic Variables 90
Contents

Data 90

Selection of Statistical Technique 92

Linear Probability Model 94

Binary Probit and Logit Probability Models 96

Empirical Results 99

Interpreting the Empirical Results 103

Comparing Ordinary Least Squares and Logit Results 107

Income 108

Traditional Noneconomic Variables 109

Special Noneconomic Variables 112

Goodness of Fit 113

Chapter 5

Demand for Residential Main Stations 115

Theoretical Models 116

Phone versus No Phone 119

Income 120

Prices 120

Noneconomic Variables 123

Omitted Variables 125

Selection of the Statistical Model 126

One-Party versus Multi-Party Service 127

Two-Party versus Four-Party Service 130

Data 132

Empirical Results 135

Phone versus No Phone 144

Income 144

Price 145

Noneconomic Variables 147

One-Party versus Multi-Party and Two-Party versus Four-Party Service 148

Goodness of Fit 150

Chapter 6

Summary and Conclusions 151

Summary 153

Economic Variables 153

Traditional Noneconomic Variables 154

Special Noneconomic Variables 156

Policy Implications 157

Implications for Regulation 157

Implications for Telephone Company Planning 163

Suggestions for Further Research 164

References 167

Appendices 175

Notes 195
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Types of Engel Curves Considered for the Toll Demand Model</td>
<td>43</td>
</tr>
<tr>
<td>2</td>
<td>Definition of the Variables Used in the Toll Demand Model</td>
<td>62</td>
</tr>
<tr>
<td>3</td>
<td>Regression Results for the Toll Demand Model</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>Sample Means and Standard Deviations for the Explanatory Variables Used in Toll Demand Model 1, Table 3</td>
<td>66</td>
</tr>
<tr>
<td>5</td>
<td>Antilogarithm of the Coefficient of Each Dummy and Household Composition Variable Used in Toll Demand Model 1, Table 3</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>Simple Correlation Coefficient Matrix for the Explanatory Variables Used in Toll Demand Model 1, Table 3</td>
<td>68</td>
</tr>
</tbody>
</table>
List of Tables

Table

7 Regression Results for the Toll Demand Model Using an Alternative Specification for Local Service Use

8 Regression Results for the Toll Demand Model Using an Alternative Specification for Geographical Dispersion of Family

9 Ordinary Least Squares Regression Results for the Extension Telephone Demand Model

10 Binary Logit Results for the Extension Telephone Demand Model

11 Sample Means and Standard Deviations for the Explanatory Variables Used in Extension Demand Model 1, Table 10, and in Table 9

12 Matrix of Simple Correlation Coefficients for the Explanatory Variables Used in Extension Demand Model 1, Table 10, and in Table 9

13 Carolina Telephone and Telegraph Company Schedule of Base Rates for Residential Main Station Telephone Service

14 Carolina Telephone and Telegraph Company Mileage Zone Charges

15 Definition of the Explanatory Variables Used in the Main Station Demand Model

16 Ordinary Least Squares Regression Results for the Main Station Demand Models

17 Binary Logit Results for the Main Station Demand Models

18 Sample Means and Standard Deviations for the Explanatory Variables Used in the Main Station Demand Models of Tables 16 and 17

19 Simple Correlation Coefficients Between the Explanatory Variables Used in the Main Station Demand Models of Tables 16 and 17

Page

80

83

100

101

103

104

121

122

128

136

138

141

143

Acknowledgments

The author expresses his sincere appreciation to Dr. R. E. Sylla for his guidance and encouragement in this study and throughout his entire Ph.D. program. The comments and suggestions of the other members of his advisory committee — Dr. E. W. Erickson, Dr. F. G. Gesbrecht, and Dr. J. O. Williams — are also greatly appreciated. The author is grateful for the financial support provided by a doctoral fellowship from the Institute of Public Utilities at Michigan State University. The author would like to thank Dr. D. M. Hoover for permission to use the data gathered in the Second Survey of Labor Supply in Eastern North Carolina and for obtaining information during the interviews to be used specifically in analyzing the demand for residential telephone services. Appreciation also goes to the Carolina Telephone and Telegraph Company, especially T. P. Williamson, W. C. Morris, B. S. Wilder, and Wendell Williams for making available telephone records, without which this study would not have been possible. The cooperation and assistance of past chairman M. R. Wooten, current chairman R. K. Koger, and members of the staff of the North Carolina Utilities Commission in providing other necessary data are also gratefully acknowledged. Special gratitude is expressed to Ms. Harriet McLaughlin for her patience in writing the numerous computer programs necessary for the empirical analysis of this study, to Ms. Ruth Good for editing the manuscript, and to Ms. Peggy Brantley for typing the final draft.
Introduction

Objectives

The purpose of this dissertation is to use cross-section data on individual households to determine which economic factors, such as price and income, which traditional noneconomic variables, such as family size and composition, urbanization, race and social class, and which special noneconomic variables, such as geographical dispersion, of family, size of housing unit and calling scope, affect the demand for several types of residential telephone service. The specific objectives of this study are to determine empirically the effect of these factors on:

1) A household's decision to subscribe to telephone service (i.e., rent a main station).\(^1\)
Chapter 1

(2) A residential telephone subscriber's choice among available grades of main station telephone service (i.e., one-party, two-party, or four-party),

(3) A residential telephone subscriber's decision to rent an extension station,

(4) A residential telephone subscriber's expenditures on long-distance or toll massage telephone service.

Usefulness of Results

The main concern in this dissertation is obtaining empirical estimates of the effects of the above-mentioned factors on the demand for residential telephone services. The results of this econometric demand study will be of value to scholars, telephone companies, and federal and state regulatory authorities.

Addition to Empirical Demand Knowledge

Recent advances in statistical techniques, together with increasingly available household budget data and market statistics, have made possible empirical studies of the demand for many different types of goods and services. However, due to the lack of necessary data, there exists little empirical literature on the demand for residential telephone services. The results of this study not only increase the understanding of the effects of factors studied in the past but also represent a significant contribution to what is known about the effects of factors not previously examined, such as family size and composition, degree of urbanization, race, social class, geographical dispersion of family, calling scope, levels of living quarters, etc.

Improved Planning and Marketing by Telephone Companies

Apart from academic interest, the findings of this study will be of use to telephone company planners and marketing researchers. Accurate forecasts of future demand for residential telephone services are a prerequisite for telephone company planning. Current demand forecasts establish expected future needs for telephone instruments, wire centers, central office equipment, and outside plant equipment necessary to tie telephone users to the switching equipment. These future capital stock requirements in turn determine the level of investment necessary to ensure that the additional facilities will become operational just as they are needed. Underestimating future demand may create a shortage in the supply of telephone services, and overestimating future demand will result in a premature expansion of facilities which will become an added cost to the company.

Currently, future demand for residential main telephones, extension telephones, and toll service is frequently estimated by extrapolating existing trends in consumption using statistical time-series analysis. Models containing economic variables, such as those developed and estimated in this research, can be used to explain and predict the growth of these services more satisfactorily than can be done with linear trend analysis. The empirical estimates obtained in this study, when combined with projections regarding future price, income, household characteristics, etc., can have significant implications regarding future demand.

Telephone company marketing researchers will also find the results of this study useful. An understanding of the characteristics of consumers of telephone services would be helpful in selecting advertising media and appeals compatible with these attributes. An identification of potential customers, made possible with the models developed here, would permit an orientation of sales and promotional campaigns to alter the responses of these groups. If the character-
istics of a given market area known, the empirical results obtained in this study can be used to estimate the sales potential of the area. Such marketing policies cannot be formulated using the results of currently available empirical studies.

Improved Regulation of Telephone Companies

The results of this study could also improve regulation of telephone companies by state and federal authorities. State utility commissions have the legal responsibility for setting local service and interstate toll service rates that telephone companies operating within the state are permitted to charge. Interstate toll rates are regulated by the Federal Communications Commission (hereafter referred to as FCC).

State utility commissions attempt to set rates so as to provide telephone companies with sufficient revenue to cover their operating expenses and to produce a fair profit for their stockholders. When the rate of return earned by a telephone company falls below what it considers to be a fair rate of return, the company may request a general rate increase that could result in authorization to raise its rates.

Rate hearings have become more frequent in recent years as a result of increases in the rate of inflation. Jacoby (1976, p. 24) reports that residential main station rates may increase drastically in the future as a result of recent FCC rulings permitting competing companies to supply equipment used on customer premises and to provide interstate private line circuits. Officials of the American Telephone and Telegraph Company (hereafter referred to as AT&T) predict that, as this competition eats into the telephone company's revenues, they will have to increase residential local service rates in order to make up for revenue lost to competitors.

The tacit assumption made in most rate cases, when determining by how much to increase prices in order to provide the authorized additional revenue, is that the demand for residential main station telephones is perfectly inelastic for a reasonable change in price. For example, in a recent rate application filed with the North Carolina Utilities Commission (hereafter referred to as NCUC) by Carolina Telephone and Telegraph Company (hereafter referred to as CT&T), estimated revenues from proposed residential main station rate increases were calculated by simply multiplying the new rate for each grade of service by the number of subscribers currently renting each grade and summing. In other words, it was assumed that an increase in telephone rates would not result in any subscribers disconnecting their telephones or changing their grade of service. If these assumptions are not valid, the price change will not provide the anticipated amount of revenue. The validity of these two hypotheses is tested in this study.

As Lowry emphasizes in a review of the literature on the demand for telecommunications services, the role of demand analysis in rate hearings has become increasingly important. Regulatory commissions now recognize the need for quantitative support for pricing telecommunications services and have begun asking for estimates of the price elasticity of demand for various services. It is hoped that the results of this study can add to the knowledge concerning these elasticities and be of use in rate hearings.

As Mitchell's study (1976) demonstrates, empirical estimates of the demand for residential main stations can also be used to analyze the effects of converting from flat rate pricing of local exchange service to "lifeline" or "usage-sensitive" pricing. These new pricing schemes, which are intended to benefit aged and low income persons, offer a lower monthly rental than flat rate service but one must pay
Chapter 1

for each local call above some minimum number. The empirical results obtained in this research can be used to analyze the effects of many other regulatory policy changes, such as expansion of base rate areas, elimination of two- and four-party service, changes in mileage zone charges, and increasing the size of calling scopes due to Extended Area Service. Available empirical evidence on residential main station demand is both limited in quantity and poorly suited for providing answers to such policy changes. The models formulated and estimated in this dissertation are capable of analyzing such policy issues.

Weaknesses of Prior Research

The limitations of previous empirical studies of the demand for residential telephone services lie in the inadequacy of the data analyzed. National, state, telephone companies, and individual telephone exchanges have been used as the unit of observation by prior researchers. Due to the unavailability of aggregate data on the quantity consumed of some particular type of residential telephone service, price-deflated revenue has often been used as the measure of consumption. Since local service revenue and toll service revenue are not usually further separated into that derived from residential subscribers and that obtained from business customers, many researchers have been unable to estimate a separate demand function for each of these four revenue categories. Previous researchers have also experienced difficulty obtaining accurate data on explanatory variables corresponding to the same unit of observation as the dependent variable. As a result, proxy variables have frequently been used. Handicapped by the unavailability of micro-level data, existing studies are able to provide only limited insight into the effects of household characteristics on the demand for residential telephone services.

Introduction

In the research reported here, the availability of extensive and relevant data on individual households allows many of the difficulties encountered in previous studies of telephone demand to be avoided. This study represents the first comprehensive analysis of the demand for all types of residential telephone service based on data from a large sample of individual families.

Scope and Limits of Study

This dissertation is an econometric study of the demand for several types of residential telephone service. The demand for telephone services by businesses is not studied. Also not considered here are elements of cost or supply and optimal pricing policies. Costs of communications services and their pricing are very complex subjects and warrant much additional study.

The data base analyzed in this research permits the effects of some but not all prices to be studied. Due to cross-sectional variation in the monthly rental rates for main telephones, it is possible to analyze the effect of these prices on a household's decision to rent a main telephone and on its choice of grade of service. However, since the rental rate for an extension telephone and the rate structure for long-distance calls are the same for all households in the sample, it is not possible to study the effect of price on the demand for these two services.

Since the empirical results obtained in this research are based on data for a sample of households in eastern North Carolina, the estimated effects of income, prices, and other explanatory variables on the demand for residential telephone services may not be applicable to populations in different geographic areas or at a different time period. Domencich and McFadden (1975, pp. 4-5) point out, however,
that if the model estimated is truly behavioral, i.e., if it adequately represents the causal relationship between purchasing decisions and household characteristics, then its estimated parameters should reflect the motivations of people in general, rather than merely the characteristics of the area from which the data used to calibrate the model were drawn. If the demand functions developed in this dissertation include all the important factors affecting the consumption of telephone services, the estimated parameters may be applicable to different situations in different areas of the country.

Description of Study

The set of data examined in this research was created by supplementing information gathered in the Second Survey of Labor Supply in Eastern North Carolina with data supplied by the telephone company providing service to this area. From mid-March through early June 1975, the Department of Economics and Business of North Carolina State University obtained completed interviews from a random stratified sample of 2,113 households living in eight eastern North Carolina counties. The counties surveyed were Beaufort, Bertie, Hertford, Martin, Pitt, Greene, Lenoir and Jones. During these personal interviews information was obtained on such general household characteristics as income, family size and composition, degree of urbanization, race, and education and occupation of the family members. The respondents also supplied information to be used specifically in analyzing the demand for telephone services, such as whether or not they rent a residential main telephone, the size of their housing unit and the location of immediate family relatives living outside their local calling area.

For each household giving its written permission, CT&T provided detailed information on grade of main station service received, number of extension telephones rented, and expenditures on long-distance calls. Additional variables representing telephone system characteristics, such as price of each grade of main station telephone service available and calling scope, were constructed from information on file with the NCUC.

The study proceeds as follows. A review of the relevant literature on the demand for residential telephone services is presented in Chapter 2. In Chapter 3 demand theory and previous studies are used to develop a model of the demand for residential long-distance telephone service. Data on a subset of the households interviewed are used to estimate a double-logarithmic demand function using ordinary least squares regression analysis.

The demand for residential extension telephones is analyzed in Chapter 4. A residential customer's decision to rent an extension station is hypothesized to be a function of the socioeconomic characteristics of the household. Since the dependent variable in this case is either zero or unity, according to whether a household rents or does not rent an extension telephone, ordinary least squares, with its assumption of a continuous dependent variable, is no longer the appropriate analytical tool. Thus, binary logic analysis, which hypothesizes a sigmoid-shaped relationship between the probability of renting an extension telephone and a linear combination of independent variables, is used to estimate the probability choice model using data on 1,018 households.

The analysis is extended in Chapter 5 to consider the factors affecting a household's decision to subscribe to telephone service and its selection of grade of service. A series of dichotomous choice models are constructed to simulate the decision-making process by those households having more than two grades of service available.
Binary logit analysis is the primary statistical technique employed in estimating each of the probability choice models formulated in this chapter.

The final chapter, Chapter 6, summarizes the results of the study, discusses policy implications of the findings, and offers suggestions for additional research.

2

Review of Previous Studies

This chapter contains a survey and critique of the existing econometric literature on the demand for residential telephone services. There are three purposes for summarizing and evaluating previous empirical demand studies. First, such a review provides the reader with a comprehensive survey of the state of the art in this area and an understanding of the difficulties encountered in earlier work. Secondly, the results of previous investigations provide a guide to the selection of explanatory variables to be included in the demand functions developed in the following chapters. Thirdly, the results of earlier studies provide a basis for judging the reasonableness of the estimates obtained in this dissertation.
Two groups of previous studies of the demand for telephone services are not reviewed in this survey. The first group not discussed consists of statistical studies in which future demand for telephone services is forecast using statistical time-series analysis to extrapolate historical trends in consumption. In these studies the dependent variable is hypothesized to be a function only of lagged values of the dependent variable. Economic and demographic variables, which demand theory suggests are important factors affecting the demand for a service, are not included in these models.

The second group of studies not considered in this review of the literature consists of studies that investigate the effects of economic variables, such as income, price and population, but are unable to analyze the demand for telephone services by households separately from business demand. In other words, the dependent variable in these studies is the sum of residential and business demand. Researchers analyzing aggregate data for the United States have been unable to separate business from residential demand largely because telephone companies, in their annual reports filed with the FCC, are not required to separate revenue obtained from business users from revenue collected from residential users and to further distinguish within these categories between local service and toll service revenue. Clearly, however, the factors affecting the demand for telephone services by households and businesses are sufficiently different that they should be analyzed separately.

The only studies described in detail in this survey are those that analyze the effects of economic variables on the demand for residential telephone services. The first five studies reviewed analyze aggregate data and the last two examine data on individual households. The chapter concludes with a discussion of the disadvantages of analyzing aggregate data and how these problems are avoided in the research reported here.

**Stern**

Stern (1965) analyzed the effects of seven separate rate increases from 1950-1959 on the growth of residential main stations for the Chesapeake and Potomac Telephone Company of West Virginia. Stern simply marked off where the rate changes occurred on a graph showing the number of residential main stations rented each month. On the basis of this crude analysis, Stern concluded that the price elasticity of demand for residential main stations approaches zero.

Stern employed a similar procedure to analyze the effects of four changes in residential extension telephone rates during this period on the trend in growth of residence extension telephones. His analysis suggests that residential extension telephones are subject to a somewhat higher price elasticity than residential main stations. It was estimated to about -.1.

The major drawback of Stern's study is that other factors, such as income, population and calling scope, which affect the demand for residential main stations and extensions, were not held constant while the effects of price were being investigated. In addition, no explicit statistical technique, such as regression analysis or analysis of variance, was employed in the research methodology.

**Dobell et al.**

As part of an econometric model of the demand sector of the telecommunications industry in Canada, Dobell et al. (1972) estimated the demand for residential local and toll service using annual data for Bell Canada for the period 1950-1967. In their analysis of residential local service they employed the flow-adjustment model of
Houthakker and Taylor (1970, pp. 26-29). This dynamic demand model is used frequently in analyzing the demand for telephone services and thus will be explained in detail. The model consists of two equations. The first equation states that desired consumption in time \( t \) (\( q^*_t \)) is a linear function of income in time \( t \) (\( x_t \)) and price in time \( t \) (\( p_t \)), i.e.,

\[
q^*_t = a + bx_t + yp_t.
\]  

The second equation expresses the rate of change in consumption at time \( t \) (\( q_t \)) as being proportional to the difference between desired consumption (\( q^*_t \)) and actual consumption in time \( t \) (\( q_t \)), i.e.,

\[
q_t = \delta(q^*_t - q_t).
\]  

Following Houthakker and Taylor (1970, pp. 26-29) the estimating equation becomes:

\[
q_t = A_0 + A_1q_{t-1} + A_2(x_t + x_{t-1}) + A_3(p_t + p_{t-1}).
\]  

where: 

- \( q \) = residence local service revenues of Bell Canada per capita/local service price index for residences,
- \( x \) = personal disposable income per capita for Ontario and Quebec (the area served by Bell Canada), and
- \( p \) = local service price index for residences (monthly rate for a two-party line in cities with a population of over 250,000).

When equation (3) was estimated using ordinary least squares regression analysis, the authors report that the sign of price was correct, but that the coefficients of all the other variables except \( q_{t-1} \) were not significantly different from zero at the 10 percent level. When price was tried by itself, its sign became positive but not significant. Consequently, price was excluded from the demand equation and the following results were obtained (\( t \)-ratios are in parentheses):

\[
q_t = -9.39 + 0.9494q_{t-1} + 0.000424(x_t + x_{t-1})
\]

\[(-0.77)\]  \[20.26\]  \[1.47\]

\( R^2 = 0.999. \)

Income elasticities of demand, calculated at the 1967 level for the variables, are .47 for the short run and 2.38 for the long run.

One difficulty with this estimation is that the dependent variable was obtained by dividing residential local service revenue by a price index for residential main stations. Price-deflated residential local service revenue may be a poor proxy for the number of residential main stations since local service revenue presumably includes revenue from extension telephones and vertical services, such as extra bells and gongs, color telephones, fancy telephones, etc., as well as from main telephones.

To estimate the demand for residential long-distance telephone service, Dobell et al. employed the stock-adjustment model of Houthakker and Taylor (1970, pp. 9-26). This model consists of two equations: a behavioral relationship stating that consumption in period \( t \) (\( q_t \)) is a linear function of income in period \( t \) (\( x_t \)), price in period \( t \) (\( p_t \)) and the period \( t \) stock of this good remaining from past consumption expenditures (\( s \)), i.e.,

\[
q_t = a + bx_t + yp_t.
\]

and a relationship expressing the rate of change in the stock at time \( t \) (\( s'_t \)) as a function of \( q_t \) and the rate of depreciation (6), i.e.,

\[
sg'_t = q_t - s_t.
\]  

In the case of services, stocks are interpreted to represent habit formation. In such a case \( \delta \) in equation (5) is expected to be positive.
The estimating equation is obtained after using equation (6) to eliminate the unobservable \( u_t \) from equation (5) and translating to discrete time. The results of estimating this equation using ordinary least squares regression are as follows (t-ratios are in parentheses):

\[
q_t = 0.24 + 0.8674q_{t-1} + 0.001176x_t + 0.00103x_{t-1} + 0.02481p_t - 0.00927p_{t-1} \\
(0.14) \quad (2.77) \quad (1.02) \quad (0.65) \quad (-4.04) \quad (-1.01)
\]

\[\hat{R}^2 = 0.993,\]

where:
- \( q \) = residential long-distance revenues for Bell Canada per capita/long-distance price index for residences,
- \( x \) = personal disposable income per capita for Ontario and Quebec,
- \( p \) = long-distance price index for residences (rate for a 3-minute, 350-mile, nighttime, station-to-station toll call).

Although the coefficients of both the income variables and the coefficient of one of the price variables are not significantly different from zero at the 10 percent level, the authors report the following elasticities, calculated at the 1967 levels for the variables:

<table>
<thead>
<tr>
<th></th>
<th>Income</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short run</td>
<td>0.20</td>
<td>-0.30</td>
</tr>
<tr>
<td>Long run</td>
<td>1.27</td>
<td>-1.90</td>
</tr>
</tbody>
</table>

Due to the large differences between short-run and long-run income and price elasticities, Dobell et al. conclude that the demand for residential toll service, as was the case with residential local service, is characterized by strong habit formation.

Review of Previous Studies

Waverman believes the demand functions estimated by Dobell et al. (1972) are misspecified and therefore the calculated price and income elasticities are incorrect. Two deficiencies in the Dobell et al. study led to this conclusion. First, little explanatory power is yielded by variables such as income and price. Secondly, the coefficient on the lagged dependent variable is unreasonably high; .95 for residential local service and .87 for residential toll service. Lagged coefficients of this magnitude imply that it would take consumers 15 to 20 years to adjust fully to an exogenous change in income or price. This is an unreasonably long lag for a service which is consumed frequently and in the home. Waverman reformulated both the residential local service and the toll service demand functions and re-estimated them with essentially the same data used by Dobell et al. 10

In reformulating the residential toll service demand model, Waverman specified the dependent variable as residential toll revenue of Bell Canada per residential main station deflated by a price index for residential toll calls. An exponential Koyck adjustment model was found to provide a better fit than either a stock-adjustment or a linear Koyck model. The estimating equation for the exponential Koyck adjustment model is given as follows:

\[
\ln q_t = a_0 + a_1 \ln q_{t-1} + a_2 \ln x_t + a_3 \ln p_t \tag{8}
\]

where:
- \( q \) = residential long-distance revenue for Bell Canada per residential main telephone/long-distance price index for residences,
- \( x \) = personal disposable income per household, and
- \( p \) = long-distance price index for residences.

When equation (8) was estimated using ordinary least squares
regression analysis, the coefficient of the lagged dependent variable was found not to be significantly different from zero. This implies, in contrast to the Dobell et al. findings, that consumers adjust their demand for toll message telephone service within one year of a change in an exogenous variable. Both income and price had coefficients significantly different from zero at the 5 percent level.

The results of re-estimating this equation with the lagged dependent variable deleted are as follows (t-ratios are in parentheses):

\[ \ln q_t = -6.057 + 1.114 \ln x_t - .970 \ln p_t \]

Since equation (9) is double-logarithmic, the estimated parameters represent elasticities. Hence, the results show an income elasticity of demand for toll calls of 1.11 and a price elasticity of -.97. Since time-series data were used, these elasticities can be interpreted as short-run elasticities.

Waverman also reformulated the Dobell et al. residential local service demand function using a different specification for the dependent variable and a different functional form. Realizing that deflating residential local service revenue (which includes revenue from extension telephones and vertical services) by a price index for residential main stations leads to a biased estimate of the number of telephones, Waverman used the actual number of residential main telephones for Bell Canada as the dependent variable in his model. Independent variables included income, price, and the number of households. A simple linear demand function was estimated using ordinary least squares regression analysis. The high degree of serial correlation between the error terms, a difficulty often encountered in analyzing time-series data, was corrected for using a search procedure developed by Hildreth and Lu (1960). The empirical results obtained are as follows (t-ratios are in parentheses):

\[ q_t = -.2432 + .0010 h_t + .502 x_t - .3107 p_t \]

where: \( q_t = \) number of Bell Canada residential main telephones,
\( h_t = \) number of households,
\( x_t = \) personal disposable income, and
\( p_t = \) price index for residential main stations (monthly rate for a two-party line in cities with a population of over 250,000).

As revealed by the t-ratios, the coefficients of all the explanatory variables included in equation (10) are significantly different from zero at the 10 percent level. The income elasticity of demand for residential main telephones, calculated at the mean of the variables, is .15 and the price elasticity is -.12.

The Waverman and Dobell et al. studies have several deficiencies in common as a result of examining the same data set. One problem with these two studies is that price-deflated revenue is often used as the dependent variable. This is an imperfect measure of quantity. Also, only a few explanatory variables could be studied. No noneconomic variables were investigated. Another drawback of these studies is that the data set consisted of only 17 annual observations. The difficulties encountered in analyzing aggregate data in general are summarized in the concluding section of this chapter.

**Long-Distance Interstate Message Toll Service Demand Model**

The American Telephone and Telegraph Company has developed a computerized mathematical model designed to evaluate the effects of proposed changes in interstate toll rates on the number of interstate
toll messages and toll revenue. This very complex model consists of two parts. The purpose of the first part of the model is to predict the number of each possible type of toll call that will be made. There are 6,750 different types or classifications of toll calls. This is the number of possible combinations of the following characteristics of long-distance calls: type of customer (business, residence, public), type of call (station-to-station or person-to-person), length-of-haul (1-8 miles, 9-13 miles, and so on up to the longest classification of 2301-3000 miles), type of day (weekday, Saturday, Sunday), and time of day (23 one-hour periods and 2 half-hour periods). Thus, a separate model was developed to explain the factors affecting the number of toll calls made in each of the 6,750 subdivisions or sectors of the interstate toll market.

The message volume in each of these sectors is hypothesized to be a function of economic conditions (ratio of disposable personal income to disposable personal income trend for residential toll calls), rate for that type of call, rates for substitute-type calls, the market's growth rate, diversion to competitive services such as Wide Area Telecommunications service and private line service, shocks (such as floods, stock market fluctuations, etc.), military conditions, etc. The exact functional form for the equation used to model the message volume of a sector for October of year 1 is as follows:

\[ M_1 = A_0 X_1 X_2 X_3 X_4 X_5 X_6 X_7 + a_1 X_1 + a_2 X_2 + a_3 X_3 \]

where:
- \( M_1 \) = number of interstate toll messages in this sector,
- \( A_0 \) = messages in this sector of the market during a base period (October 1960),
- \( X_1 \) = military variable,
- \( X_2 \) = economic variable,
- \( X_3 \) = price variable,
- \( X_4 \) = competition and shocks variable,
- \( X_5 \) = year associated with growth rate prior to 1960,
- \( A_6 \) = growth rate for 1948-1960,
- \( X_7 \) = year associated with the growth rate after 1960,
- \( A_8 \) = growth rate for 1960 to present,
- \( X_9 \) = equivalent number of business days,
- \( X_{60C1} \) = number of messages shifted from cell \( C_1 \), and
- \( a_1, a_2, a_3 \) are parameters to be estimated.

The additive term at the end of equation (11) is designed to take account of potential shifting from one sector of the market to another due to changes in the rates of substitute sectors.

The parameters of equation (11) were estimated for each of the 6,750 sectors by applying a combination of ordinary least squares regression analysis and successive approximation to data for the United States from 1948 through 1965. The results of this analysis indicate that the residential demand for toll calls is price inelastic, with the elasticity increasing with the length of haul. Although actual income elasticities were not calculated, the estimated coefficients of \( a_2 \) for residential toll calls suggest that the income elasticity of demand for interstate toll calls increases with the length of haul.

In order to determine the effects of rate changes on revenue, it is necessary to know the length of conversation for each toll message. The second part of the AT&T model consists of 6,750 equations representing the distribution of the length of conversation for each sector. The average number of overtime periods can then be calculated. When this figure is multiplied by the overtime rate and added to the initial period rate, the average revenue per message in each sector can
be derived. When this figure is multiplied by the predicted number of messages in each sector, an estimate of total revenue for each sector is obtained.

With the data used in estimating the parameters of this model, it would be possible to calculate an income elasticity of demand for interstate toll calls for residential subscribers. Such an estimate, which would be very useful for comparing with the income elasticity calculated in this dissertation, is unavailable.

Irish

The only study analyzing the demand for particular grades of residential main stations and residential extension telephones is that of Irish. Pooled cross-section and time-series data for 1968, 1970, and 1972 for individual telephone exchanges in North Carolina were examined. For each exchange, data on the number of residential telephones of each grade, the number of residential extensions, and calling scope were obtained from station development reports filed by North Carolina telephone companies with the North Carolina Utilities Commission. Base rates for each grade, installation charges, and monthly rental rates for extensions were obtained from tariff sheets also filed with the NCUC.

Unfortunately, proxy variables had to be used for many of the explanatory variables included in the demand functions. For example, the population of the town or city corresponding to each exchange was used as an estimate of the population of the exchange and, in some models, as an estimate of the base rate area population. Per capita retail sales for the county in which each exchange was located were the best measure of income available.

Some of the results obtained by Irish will now be discussed.

Demand for a particular grade of service was hypothesized to be a function of the monthly base rate for that grade, monthly base rate for the closest substitute grade, income, population, calling scope, and whether the exchange is a Southern Bell exchange or an exchange of an independent telephone company.

Two different specifications of the dependent variable were used in estimating the demand for one- and two-party main stations. The first specification was the total number of residential telephones of each grade in the entire exchange. The problem with this specification is that the price variable corresponds to the monthly rental rate for local service within the base rate area of an exchange where mileage charges do not apply. To adjust quantities to correspond with base rate area prices, an estimate of the number of one-party main stations within the base rate area of each exchange was obtained by multiplying the number of one-party main stations in the whole exchange by the ratio of residential private line telephones in the base rate area to total residential one-party telephones for the exchange for November 1967. A similar procedure was used to estimate the number of two-party main stations within the base rate area. Only the results obtained with this alternative specification are discussed.

Both linear and double-logarithmic functional forms were employed in estimating the demand for one-party main stations within the base rate area. One of several double-log demand functions estimated using ordinary least square regression analysis is presented below (the probabilities of a greater absolute value of t are in parentheses):

$$\ln RMTIBRA = -4.053 + 3.303 \ln PR1 + 2.546 \ln PR2 + .862 \ln PE + .661 \ln RSPC + .026 \ln CS + .829 \ln MD$$

$$R^2 = .722,$$

(12)
Chapter 2

where:  
RMTRA = estimate of the number of one-party residential main stations within the base rate area of an exchange,  
PR1 = residential one-party monthly base rate,  
PR2 = residential two-party monthly base rate,  
PE = population of the town or city in which the exchange is located,  
RSPC = per capita retail sales for the county in which the exchange is located,  
CS = calling scope of the exchange, and  
D = 1, if Southern Bell exchange; 0, if otherwise.

The coefficients of the two price variables, population, and income in equation (12) are significantly different from zero at the 1 percent level and have the expected signs. The elasticities for each of these variables are given by the coefficients. Calling scope does not exert a significant effect on the demand for one-party main stations.

Linear and double-log functional forms were also used to estimate the demand for two-party residential main stations within the base rate area (RMT2BRA). The results of estimating the latter demand function using ordinary least squares regression are as follows (the probabilities of a greater absolute value of t are in parentheses):

$\ln RMT2BRA = -1.853 + \frac{.942}{\ln PR1} - \frac{1.850}{\ln PR2} + \frac{.661}{\ln PE} + \frac{.387}{\ln RSPC} + 0.0356 \ln CS$

(13)  

$R^2 = .581,$

where all the variables are as defined previously. Although the coefficients of all the explanatory variables in equation (13) have the theoretically anticipated sign, price of one-party service and calling scope do not exert a significant effect on the demand for residential two-party main stations.

Very poor results were obtained in the attempt to estimate a demand function for four-party residential main stations. This probably stems from the fact that the dependent variable in this model refers to the number of four-party residential main stations in the whole exchange, whereas prices apply to the base rate area, within which four-party service is not even available in many exchanges.

In analyzing the demand for residential extension telephones, Irish did not have to adjust quantities to correspond to prices. The monthly rental rate for an extension telephone is the same in all sections of an exchange. The results of estimating one of several double-log demand functions are presented below (probabilities of a greater absolute value of t are in parentheses):

$\ln RET = -1.278 - .511 \ln PRE - .267 \ln PR1 + .131 \ln RET + \frac{1.037}{\ln RI} + 0.107 \ln RSPC + 0.276 \ln QSI$

(14)  

$R^2 = .913,$

where:  
RET = number of residential extension stations in an exchange,  
PRE = monthly rental rate for a residence extension telephone,  
RI = cost of installing an extension station,  
RI = number of one-party residential main stations in an exchange,  
QSI = proxy for quality of service (ratio of one-party main stations to total number of main stations in the exchange), and  
PR1 and RSPC are as defined previously.

The coefficients of all the explanatory variables in equation (14) are significantly different from zero at the 10 percent level, and all but installation cost have the correct sign. These results indicate that the income elasticity of demand for residential extensions is .107 and the price elasticity is -.51.

The primary drawback of Irish's study is that proxy variables had to be relied on to measure many of the important variables included in
the demand functions. The accuracy with which these variables, particularly income, were measured is questionable.

The approach employed by Irish to study the demand for residential main stations and extension stations is different from that taken in this dissertation. Irish attempted to estimate market demand functions for each grade of service and extension telephones, i.e., to determine the factors affecting the total number of 1-, 2- and 4-party main stations and extension stations in an exchange. In this dissertation micro-level data are analyzed first to determine the factors affecting a household's decision to subscribe to telephone service and then to determine the factors affecting its choice among available grades of service and its decision to rent an extension telephone. Thus, Irish's estimates are not directly comparable with those obtained in this study.

Perl

Perl conducted a very thorough analysis of the economic and demographic factors that influence whether households in the United States have a telephone available at which they can be reached. Perl realizes that some households with a telephone available do not subscribe to telephone service. The effects of income, telephone rates, urbanization, education, race, region, and family type (i.e., stage of the life cycle) on telephone availability were examined.

Data on all the variables except telephone rates were obtained from the 1/1,000 Public Use Sample of the 1970 Census of Population conducted by the U.S. Bureau of the Census (1972b). A price for telephone service was assigned to each of these 36,671 households according to the following procedure. First, the location of each household in the Census sample was identified by county or county aggregate. Each of these county areas was assigned to one of 100 Bell System Revenue Accounting Offices. Data from AT&T's Market Research Information System and ordinary least squares regression analysis were used to estimate the average rate paid by the minimum income family in each Revenue Accounting Office. This price was assumed to reflect the rate for the lowest grade of residential main station service available. Each household was then assigned the price corresponding to the Revenue Accounting Office in which it was located.

There are three difficulties with this price variable. First, all households assigned to the same Revenue Accounting Office will be assigned the same price, although those families living a greater distance from the center of their exchange, and thus subject to mileage zone charges, or living in an exchange with a larger calling scope will actually be faced with a higher rate for telephone service than the one attributed to it by Perl's procedure. The second problem is that the grade of service corresponding to the minimum price for basic telephone service may not even be available to all households to which this price has been assigned. This would be especially true for households located within the base rate area of an exchange where four-party service is obsolete, as this service is in many areas of the country. The third difficulty with this price variable is that the lowest price at which telephone service is available may not be the price affecting the decision to rent a telephone by households having two or three grades of service available.

Two analytical techniques were used by Perl to study these data. First, the data were cross-tabulated by price, income, and selected household characteristics. Two, 3, and 4-way frequency tables were constructed to show the percentage of households with various characteristics having a telephone available. In addition, ordinary least squares regression analysis was employed to estimate a linear
Chapter 2

relationship between the probability of a household having a telephone available and a linear combination of economic and demographic variables.

Perl's extensive analysis of the data produced many interesting results. The probability of a household having a telephone available was found to increase with income, age, and education of the family head. Urban households were found to be more likely to have telephones than those living in rural areas, ceteris paribus. Whites were discovered to be more likely to have telephones than blacks. Families on welfare were found to be less likely to report a telephone available than similarly situated households not on welfare. The results also show that male individuals and other male-headed families are less likely to have a telephone available than other family types. Perl's cross-tabulations clearly indicate that the effects of many of these variables on telephone availability are nonlinear.

The estimated minimum price for telephone service was found to exert a significant negative effect on telephone availability. Furthermore, the effect of price was found to depend on family income as well as on the age of the head of the household and the type of family. Price elasticity of demand was discovered to decrease as income increases and as the age of the head increases. Price elasticity was found to be higher for male individuals, other male-headed households, and for female-headed households than for husband-wife families or female individuals. In general, these price elasticities were less than -0.1.

Perl's study represents the first attempt to analyze the demand for residential main stations using data on individual households. The results of this study will be helpful in formulating the main station demand models in Chapter 5 of this dissertation. The drawbacks of the Perl study may be summarized as follows: the dependent variable refers to whether a household actually rents a main station; the price variable assigned to each household may be a poor proxy for the actual price structure faced by some households; and the frequency table analysis indicates that the linear relationship postulated to exist between telephone availability and a linear combination of explanatory variables in the regression analysis may be an inappropriate specification of the functional form. These difficulties will be avoided in the empirical analysis of residential main stations reported in the present study. In addition, the analysis will be extended to consider the factors affecting choice among available grades of main station service.

Survey Research Center Study

A very interesting study was conducted by the Survey Research Center at the University of Michigan to detect which socioeconomic factors, personality factors, and attitudes toward the telephone affect long-distance calling by residential telephone subscribers. Data on a sample of 400 residential subscribers of the Michigan Bell Telephone Company were analyzed. Information on the economic, demographic, and psychological characteristics of these households was obtained in personal interviews. Data on long-distance calls placed by each of these households during a three-month period were provided by the telephone company. On the basis of these long-distance records, each subscriber was classified as either a low toll caller (no long-distance calls to places over 50 miles away during the three-month period) or a high toll caller (six or more toll calls to places 50 miles or more away).

Unfortunately, the primary analytical technique employed in the Survey Research Center study consisted of comparisons between the distribution of high toll users among categories of each explanatory
variable and the corresponding distribution for low toll users. In other words, one-way frequency tables for both groups of households were calculated for each explanatory variable and compared. The researchers recognized that this is a rather crude analytical technique since other explanatory variables cannot be held constant when each is being analyzed separately. Thus, one multiple regression equation was also estimated using ordinary least squares. The dependent variable in this equation took on a value of unity if the household was a high toll user and a value of zero if the household was classified as a low toll user. Scales, rather than dummy and quantitative variables, were used to specify each of the explanatory variables.

The Michigan study produced many interesting results, only the most important of which will be summarized here. The first group of explanatory variables analyzed consisted of various socioeconomic variables. It was discovered that people who have friends or relatives living at a distance tend to be more frequent users of long-distance telephone service than those who do not. It was also found that families with higher incomes make more long-distance calls. Due to the $(0,1)$ nature of the dependent variable used in the regression equation, it was not possible, however, to calculate an income elasticity. Young married couples without children and older married couples whose children have left home were found to be more frequent users of toll message telephone service than families at some other stage of the life cycle.

The second group of explanatory variables considered consisted of personality factors, such as security-insecurity, need for affiliation, and underlying attitudes towards the use of money. The multiple regression analysis revealed that there is no significant simple relationship between any of these personality characteristics and long-distance calling behavior.

The third group of independent variables studied consisted of proxy variables representing attitudes toward the telephone. Subscribers who use the telephone to make and receive many local calls and/or who have more than one extension telephone were found to be more frequent users of long-distance telephone service. People who make a large number of toll calls were found also to write many letters and make many visits to friends and relatives living in distant cities. The investigators discovered no tendency for either high or low toll users to be recipients of many more toll calls than they made. A larger percentage of high toll users reported experiencing a pleasant feeling while making a toll call than was the case for low toll users.

The results of the Survey Research Center study, the first to use data on individual households to analyze residential toll demand, will be very useful in formulating the long-distance demand model described in the next chapter of this dissertation. The primary disadvantage of the Michigan study lies in the analytical techniques employed. One-way frequency tables do not hold other explanatory variables constant. Only one multiple regression equation was estimated and it included only some of the explanatory variables. Furthermore, a $(0,1)$ dependent variable was used in this equation, prohibiting precise quantification of the effects of each independent variable on toll expenditures. In this dissertation more sophisticated statistical procedures will be used to analyze the factors affecting residential toll calling.

Difficulties Encountered in Analyzing Aggregate Data

Only the last two studies reviewed in this chapter employ individual households as the unit of observation. The drawbacks of these two studies have been pointed out. All the other existing econo-
metric studies of the demand for residential telephone service examine aggregate time-series or cross-section data.

The difficulties encountered in analyzing aggregate data are many. One problem hampering many attempts to use aggregate data to analyze the demand for various types of telephone service was finding a unit of observation for which accurate data existed for all the variables in the demand function. Previous researchers were forced to use proxy variables for many independent variables for which aggregate data corresponding to the same unit of observation as the quantity data did not exist. Also, due to the lack of quantity data for some classes of telephone service, earlier analysts had no alternative but to use price-deflated revenue as the dependent variable in the estimating equations. In many cases it was not even possible to separate business revenues from residential revenues.

Analysis of aggregate data also presents many statistical problems, such as multicollinearity, autocorrelation, and simultaneous equation bias. Furthermore, aggregate data cannot be used to shed any light on how noneconomic factors, such as family size and composition, urbanization, social class, race, etc., affect consumption of residential telephone services.

Additional difficulties are encountered when aggregate time-series data are analyzed. Due to the small number of observations, prices, income, and other explanatory variables may not display enough variation for their effects to be estimated. There is also the possibility of structural changes taking place over time.

Due to the good fortune of being favored with micro-level data that are more extensive and probably of higher quality than those available to previous researchers, many of the difficulties encountered in earlier studies are avoided in the research reported in this dissertation. The unit of observation used in this study is the individual household, not an aggregate of households. Data on the economic and demographic characteristics of a large sample of households living in eastern North Carolina have been supplemented with information for the same households on expenditures for toll calls, number of extension telephones and grade of main station service. These data on the consumption of telephone services were obtained directly from the telephone company providing service to this region.

The advantages of working with such data, as opposed to aggregate data, are many. Accurate information on the quantity consumed of the various types of residential telephone service is available for the same unit of observation, the household, to which the explanatory variables correspond. Thus, few proxy variables need be used. This data set also provides a unique opportunity to assess the effects on the demand for residential telephone services of household characteristics not previously studied. Cross-sectional data minimize the dangers of multicollinearity, autocorrelation, and simultaneous equation bias that characterize the analysis of aggregate data. The larger number of observations and greater variation in the independent variables permit more precise quantification of the effects of the explanatory variables than is possible with aggregate data. Cross-section data are generally regarded as producing more accurate estimates of price and income elasticities than time-series data.

Although the main objective of a demand study is predicting the behavior of aggregates of people, Domencich and McFadden (1975, p. 11) suggest that their behavior can probably best be understood by analyzing the behavior of individual consumers. Analysis of household budget data also increases the basic understanding of the decision-making process involved in consumer decisions.
3

Demand for Residential Long-Distance Telephone Service

In this chapter the demand for residential long-distance or toll message telephone service is analyzed. Demand theory and previous studies in this area are used first to determine which factors may be important determinants of a household's consumption of long-distance telephone service. This demand function is then estimated by applying ordinary least squares regression analysis to data on a subset of households interviewed in the Second Survey of Labor Supply in Eastern North Carolina. The regression results are then examined and interpreted.

Theoretical Model

The purpose of this part of the chapter is to develop a theoretical model of the demand for residents long-distance telephone service.
The unit of analysis, the dependent variable, and each of the explanatory variables hypothesized to influence a household’s usage of toll message service are discussed in turn. For each independent variable, a theoretical justification for its inclusion in the model is provided, in addition to a discussion of how it is measured and taken into account in the empirical analysis. Also, the expected direction and magnitude of effect of each explanatory variable on the consumption of toll telephone service are considered.

Unit of Analysis

Economic theory provides no identification of the consumer unit to which classical utility theory refers. It could be an individual, but it could also be a group of individuals living together. To avoid the difficult problem of assigning household income and expenditures to individual household members, most demand analysts use a household, a family, or a spending unit as the unit of observation. These approaches assume the existence of a household preference or utility function which, as David (1962, p. 6) points out, can be viewed as the end product of a set of decision rules and interactions between family members. Since all the occupants of a housing unit can reasonably be expected to make their long-distance calls from the same telephone, the appropriate unit of analysis, and the one employed in this study, is the household. The terms household and family are used interchangeably.

Dependent Variable

The behavior to be modeled is the amount of long-distance calling done by residential telephone subscribers. As in most empirical demand studies, the dependent variable in this analysis is actually an aggregate commodity. In this case the subitems are the different varieties of toll calls that result from different classes of toll calls (station-to-station or person-to-person), being placed at different times of day (day, evening, night, or weekend), and being paid for in different ways (sent paid, third party, credit card, or collect).

When analyzing the demand for an aggregate commodity, Wold and Jureen (1953, pp. 108-109) suggest defining the dependent variable in the estimating equation to be either the sum of the subitem quantities, if they are not too different, or the sum of expenditures on the subitems. Subitem quantities in this case are the number of message-minute-miles of each variety of toll call, where a message-minute-mile is one long-distance call lasting one minute and covering a distance of one mile. Expenditures, rather than message-minute-miles, are used as the dependent variable in this study for two reasons. First, regulatory authorities are more interested in knowing the effects of changes in income and other explanatory variables on expenditures, and hence telephone company revenue, rather than on message-minute-miles of toll calls. Secondly, the quantity approach would necessitate calculating the length-of-haul (i.e., distance covered) for each long-distance call placed by each household in the sample during the period of study. Expenditures on toll calls, on the other hand, were readily available.

When the dependent variable in a demand equation is an aggregate commodity, the calculated income elasticity, as noted by Wold and Jureen (1953, p. 113), is a weighted average of the income elasticities of the subitems, where the weights are the relative expenditures on the subitems. The elasticity of expenditures on toll calls with respect to income is larger than the elasticity of message-minute-miles on toll calls with respect to income if, as would be expected, an increase in income induces telephone subscribers to shift to a more...
expensive quality of toll call (such as calling during the evening instead of at night).

For those households interviewed in the Second Survey of Labor Supply in Eastern North Carolina that gave written permission, expenditures on long-distance calls (including charges for collect calls received and a 7 percent federal excise tax) for an eight-month period during 1975 were provided by Carolina Telephone and Telegraph Company. Obtaining such information directly from the telephone company produces more accurate and reliable expenditure data than could be gathered by asking respondents to record or recall their expenditures on toll calls.

In order to use data on toll charges and household characteristics for a sample of families to determine which factors affect long-distance calling, expenditures on toll calls have to be measured over a period long enough to be representative of the behavior of those families. Usually, for frequently purchased goods and services, short periods of reference provide sufficient reliability. One would not expect this to be true for toll calls since many unusual circumstances (such as births, deaths, illnesses, birthdays, etc.) affect calling behavior. Even though expenditures measured over an eight-month period should be reasonably representative of most subscribers' behavior, the sample may still contain some families demonstrating atypical calling behavior during this period.

Income

A major conclusion of classical utility theory is that the purchasing power of a household is an important factor determining its expenditure pattern. Moutakker (1952, p. 5) considers the estimation of the effect of income on expenditure to be the most important output of a demand study.

The measure of purchasing power employed in this analysis is total family gross income received during 1974. This income figure includes wages, salaries, rent, interest, and transfer payments. Two measures of this income variable were available for use in the empirical analysis. The first, reported income, is the respondent's estimate of which of eight intervals their gross family income falls. The second, an estimate of each household's total gross income, was calculated from partial information supplied on each family member's occupation, hours worked, wage rate, income from self-employment work, and transfer payments. If one of these pieces of information was missing for a particular member, a value was assigned equal to the mean for respondents with similar characteristics.

Measurement Problems. There are several measurement and theoretical problems associated with the use of this income measure. One measurement difficulty with this income figure is that it pertains to gross instead of disposable income. Since consumers probably base their expenditure decisions on income after rather than before taxes, it would be preferable to use the former income concept in the estimating equation. Disposable income, however, was not used since it would have required estimating the relevant state and federal personal income, estate, inheritance, poll, and other taxes for each family.

If the tax structure can be assumed to be progressive, using gross income instead of disposable income will result in the estimated income elasticity being biased downward. This bias, however, is not expected to be large. Analysis of 1960 - 1961 Consumer Expenditure Survey data by the U.S. Bureau of Labor Statistics (1971, p. 53) indicates that most families remain in the same income interval after tax payments are deducted.

A second measurement difficulty with the income measure used here is that payments in kind received by an employee, products taken for
home use from a family-operated farm or business, imputations for the rental value of owner-occupied homes, and savings due to performing one's own car and home repairs have not been included. Pearl (1968, p. 20), Hymans and Shapiro (1974, p. 263), and Laneing and Morgan (1971, p. 270) recommend adding these imputations to measured income so that the measure of purchasing power reflects the total flow of resources that is available to a household. The effects of omitting these imputations are difficult to predict, a priori.

Theoretical Problems. In addition to these measurement difficulties there are two theoretical problems associated with using current measured income. The approach followed in this study of regressing expenditures on current measured income and other explanatory variables implicitly assumes that individuals have had time to fully adjust their toll calling to their observed situations. In other words, this essentially static approach assumes that a household makes a change in consumption, if any, as soon as there is a change in income. Many demand analysts do not believe that this static approach provides a realistic description of how consumers actually behave. Due to habits or lags, consumers may react to an income change with some delay and their adjustment to a new equilibrium may be spread over several time periods.

This dynamic view of income has been useful in analyzing the demand for many goods and services. In their analysis of time-series data for Bell Canada, Dobell et al. (1972) concluded that the demand for residential long-distance telephone service is characterized by strong habit formation. Waverman (1973, see footnote 9) applied a more appropriate functional form to the same data and found that residential consumers adjust their demand for toll service within one year of a change in income. Unfortunately, none of the available techniques for taking account of lags and habits can be utilized in this study to shed new light on this controversy. Information on the necessary variables is lacking.

If past levels of income do affect consumption but are not included in the estimating equation, the estimated coefficient of current income will be biased upwards. This bias may not be important if most families have been experiencing a fairly stable income level. Morgan (1964) studied family income data gathered in three successive Surveys of Consumer Finances and concluded that an appreciable amount of income change does occur from year to year. Only 37 percent of the households included in the toll demand sample had their 1974 estimated gross income falling into the same category as their 1972 income. Even if income has recently changed for the families studied in this analysis, if everyone's income over the past years has changed at the same percentage rate, Currie et al. (1972, pp. 131-133) demonstrate that an income elasticity estimated from a double-log demand function, such as used here, would remain unbiased.

Actually, some allowance has been made for lags in consumers' adjustments to income changes due to the nature of the data available. Consumption refers to toll expenditures during eight months of 1975, whereas income refers to total family income received during 1974. Thus a lagged income variable is actually being used.

An additional theoretical objection to using current measured income as the index of purchasing power is that it does not adequately represent the permanent income upon which many economists contend the expenditure pattern of a household depends. According to the permanent income hypothesis, consumers do not adjust their expenditures to income received each period but instead base their purchasing decisions on what they consider to be their permanent or normal income. Thus, a consumer's spending behavior during some period depends not
only on that period's income but also on the future stream of income the consumer expects to receive.

Although much of the research on the permanent income hypothesis has been in terms of the consumption function, Friedman (1957) assumes that it can also be applied to individual goods and services. The empirical evidence on the permanent income hypothesis has not been conclusive in its support. Although several techniques are available for testing the permanent income hypothesis, they could not be employed with the data available in this study.

Using measured income instead of permanent income results in underestimates of the true income parameter. In addition, the coefficients of other explanatory variables, such as education, occupation, race, family size, etc., will be biased if they are correlated with permanent income. The extent of these biases, however, is not expected to be as large for a service such as long-distance telephone calling as it would be for durable goods.

Functional Form. In order to estimate the effect of income on toll expenditures, an explicit functional form for this relationship must be selected. The only functional forms considered in this analysis are those that can be estimated using ordinary least squares regression procedures. The mathematical forms considered include the following: linear, quadratic, inverse (hyperbolic), semi-logarithmic, double-logarithmic, log-inverse and log-log-inverse. The equation and corresponding income elasticity formula for each of these forms are presented in Table 1. For convenience, other explanatory variables are omitted from the equations. The advantages and disadvantages of each of these functional forms will now be discussed.

Although the linear form is the simplest to use, its characteristic of a constant marginal propensity to consume may result in its

---

### Table 1. Types of Engel curves considered for the toll demand model

<table>
<thead>
<tr>
<th>Form</th>
<th>Equation</th>
<th>Income elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>( E = a + bY )</td>
<td>( \eta = bY/Y )</td>
</tr>
<tr>
<td>Quadratic</td>
<td>( E = a + bY + cY^2 )</td>
<td>( \eta = (b + 2cY)/E )</td>
</tr>
<tr>
<td>Inverse</td>
<td>( E = a - b/Y )</td>
<td>( \eta = b/E )</td>
</tr>
<tr>
<td>Semi-log</td>
<td>( E = a + \ln(Y) )</td>
<td>( \eta = \partial )</td>
</tr>
<tr>
<td>Double-log</td>
<td>( \ln(E) = a + b\ln(Y) )</td>
<td>( \eta = b/E )</td>
</tr>
<tr>
<td>Log-inverse</td>
<td>( \ln(E) = a - \theta/Y )</td>
<td>( \eta = \theta/Y )</td>
</tr>
<tr>
<td>Log-log-inverse</td>
<td>( \ln(E) = a - \theta/Y + \gamma\ln(Y) )</td>
<td>( \eta = (\theta/Y) - \gamma )</td>
</tr>
</tbody>
</table>

*E* = expenditures; *Y* = income; *\( \eta \)* = income elasticity; *a*, *b*, *c*, *\theta*, and *\gamma* are estimated coefficients and \( \ln \) refers to natural logarithm.

---

diminish with increases in income. Crockett and Frieid (1960, p. 11), Pras and Houthakker (1971, p. 47) and most other demand analysts agree that a linear relationship provides a poor fit when applied to data with the range of income usually found in household budget data.

One method of allowing for a curvilinear relationship between income and toll expenditures is to specify a quadratic function. The disadvantages of this form are that income and income squared may be highly correlated or the scatter of expenditure-income points may not be large enough to determine the appropriate degree of curvature.

The semi-logarithmic form also permits a nonlinear relationship between income and expenditure. With this Engel curve both marginal propensity to consume and income elasticity vary with the level of income. A drawback of this form, pointed out by Atchison and Brown (1954, p. 36), is that income must reach some positive level before
any expenditure takes place. In their analysis of United Kingdom family budget data, Prais and Houthakker (1971) found the semi-log form to be most suited to necessities.

The double-logarithmic form also produces a curvilinear relationship. This Engel curve passes through the origin, which is considered to be disadvantageous by Prais (1953b, p. 94), and has a constant income elasticity, which Moll and Jureen (1951, p. 271) and Brown and Denton (1972, p. 1173) consider to be a drawback. However, the double-log Engel curve was found by Prais and Houthakker (1971) to give the best statistical fit for all goods and services except food. Podder (1971) found this functional form to give the best fit in an analysis of Australian budget data. As Houthakker (1965, p. 278) points out, the double-log function "remains without serious rivals in respect of goodness of fit, ease of estimation, and immediacy of interpretation."

Two other nonlinear forms, the inverse and the log-inverse, have the characteristic that consumption tends to a finite limit as income goes to infinity (i.e., there is a saturation point). Prais and Houthakker (1971, p. 82) suggest that when the dependent variable is quantity consumed the Engel curve selected should feature a satiety level. However, when consumption is measured by expenditures on an aggregate commodity where the subitems are different qualities of the good, as is the case here with toll calls, the notion of a satiety level is less applicable. One would expect the Engel curve for expenditures on long-distance calls to continue to rise in the upper income levels since subscribers may shift to higher quality toll calls as well as place more and longer calls as income increases.

The log-log-inverse form depicts a commodity as first being a luxury, then a necessity, and finally an inferior good as income increases. A priori, one would not expect long-distance calling to exhibit an inferior good phase. In addition, Brown and Denton (1972, p. 1173) found it rare for the observed range of income to be sufficient to show all three phases of some commodity.

Although all seven of these functional forms have been empirically tested, the results of which are discussed in the empirical part of this chapter, on the basis of previous studies of other goods and a priori considerations, the quadratic, double-log and semi-log are expected to provide the best approximation to the true Engel curve.

Prices

In addition to income, the classical theory of consumer behavior suggests that consumption of a particular commodity depends on the price of that commodity as well as on the prices of close substitutes and complements. Normally, when household budget data are being analyzed, prices are not included as independent variables in the demand equation since all households are presumed to face the same prices at the same point in time. Since intrastate toll rates are regulated by the NCUC and interstate rates are set by the FCC, there is complete certainty in this analysis that all households face the same toll rate structure. Also, since the sample was drawn from an eight-county area in eastern North Carolina, one would not expect geographical variation in the prices of letter writing and visiting friends and relatives in distant cities, both of which may be possible substitutes for long-distance calling. Thus, no price variables are included in the toll demand function.

Traditional Noneconomic Variables

In order to use household data to obtain an unbiased estimate of the income elasticity of demand for toll calls, it is necessary to isolate or hold constant the effects of noneconomic elements, such as
demographic, psychological, sociological, cultural and regional factors, if they are correlated with income. These factors, which affect a household’s tastes and preferences, should be included in the analysis even if they are not correlated with income in order to attain a better understanding of the variables accounting for differences in toll expenditures among subscribers.

The effects of varying tastes and preferences among households can be taken into consideration either by estimating a separate regression equation for each group of homogeneous households or by introducing additional variables representing family size, race, urbanization, social class, calling scope, geographical dispersion of family, etc., into a single equation. The latter approach was adopted in this study.

An advantage of the single equation approach is that the effect of each non-economic variable on long-distance calling can be directly determined. An important restriction inherent in this approach is that it assumes the absence of any interaction between the independent variables. In other words, it assumes that the effect of a change in one of the explanatory variables on toll expenditures does not depend on the level or value of some other explanatory variable. For example, it implies that an increase in income will have the same effect on toll charges regardless of the race, social class, or degree of urbanization of the household. The validity of this restriction is tested in the concluding section of this chapter.

Each of the non-economic variables included in the theoretical model will now be discussed. Traditional non-economic variables of the type used in most empirical demand studies are discussed first and then some specific to the service under analysis, which are referred to as special non-economic variables, will be introduced.

Household Size and Composition. Most family budget analysts agree that family size and composition (i.e., age, sex, and marital status of household members) are the most important factors affecting preferences.

As Currie (1972, p. 23) points out, a household’s consumption pattern will depend on its particular configuration of biogenic and psychogenic needs. This in turn will depend on the household’s composition, that is, on the age and sex of the members.

David (1962, p. 7) expects family composition to be associated with variation in tastes and preferences since “the impact of learning, aging, raising children, and other time-dependent processes is summarized by the family composition at any one time” and because “the expectations and planning horizons of the family are implicit in the family composition.”

As Friggle (1954) points out, difficulties encountered in taking account of household size and composition have characterized demand studies since the earliest days. The applicability of some of the methods employed in previous studies for use in modeling the demand for toll telephone service will now be briefly considered.

The three simplest methods for taking account of family size and composition (i.e., regressing consumption on family size and income, regressing consumption per capita on income per head, and regressing consumption per capita on income per capita and family size) were all rejected because, among other reasons, none of these approaches takes into consideration the age and sex of the family members. It is desirable to investigate the possibility that individual family members contribute differently to the household’s toll expenditures according to their age, sex, and other characteristics.

The best known procedure for incorporating the age and sex composition of the household into the demand function is to measure household size in terms of equivalent adults or unit consumers.
standard person usually adopted as a basis for comparison is the adult male. Other person types (e.g., persons in other age-sex categories) are expressed as fractions of the adult male; hence the term equivalent adult. With this approach the measure of family composition is the number of equivalent adults in the household. This variable is then used to deflate consumption and income. The equivalent adult scale used to deflate consumption is different from that used to deflate income, reflecting the fact that an additional person has two effects on a household's consumption of a commodity. The specific effect represents the need by that person for the commodity whereas the income effect represents the change in consumption attributable to the fact that, with a given family income and an additional person, the family has become relatively poorer. For each type of person there is a specific coefficient for each commodity. The income coefficient for each type of person is the weighted average of the specific coefficients for that type person where the weights are the relative expenditures on the commodities. Once these coefficients have been estimated, consumption, deflated by the number of household members weighted by their specific coefficients for that commodity, is expressed on income divided by the number of persons weighted by their income coefficients. Since data on expenditures on all commodities are required to estimate the income scale, this approach could not be employed. 31

A life cycle variable is another frequently used method for taking account of family size and composition. With this approach each household is placed in a life cycle category on the basis of marital status, age of head, and presence and age of children in the family. According to David (1962, p. 22), these categories

Although the life cycle approach has been found useful in analyzing consumption of some durable goods, it does not seem that the stage in their life cycle that a household is currently in would affect their toll calling. 32 Another disadvantage of this approach is that it does not provide as fine a breakdown for the number of family members and their ages as the method here adopted.

The procedure employed in this analysis to take account of both family size and composition is the one utilized by Rayner et al. (1972) and Hamid Miah (1972). The number of family members in each age-sex category is introduced into the demand equation along with family income and other explanatory variables. The number of household members in each age-sex group was calculated for the survey year, 1974, and not for the time of the interview or for the period of observation on toll calling behavior. Age was estimated for those family members not supplying this information to the interviewer. The age-sex categories employed in this analysis are as follows: children 8 years of age or younger, children between the ages of 9 and 14, adolescents between the ages of 15 and 17, adults between the ages of 18 and 21, females 22 years or older, and males 22 years of age or older.

In predicting the sign and magnitude of the coefficient of each of these household composition variables, it must be kept in mind that they are estimates of the total effect (i.e., the sum of the income effect and specific effect) that an additional person of each type will have on toll charges. 33 A priory, we would expect smaller coefficients for children and adolescents since their specific effects (i.e., relative requirements or need for toll usage service) are likely to be smaller than those of adults. 34 As long as the income effect does not outweigh the specific effect for some age-sex group, the signs of all these coefficients will be positive.

Degree of Urbanization. Another possible source of variation in
long-distance calling among households is their degree of urbanization. According to Burk (1968, p. 205) and Lee and Phillips (1971, p. 573), consumption patterns among households in different urbanization groups may differ due to differences in the following: nonmoney income; income stability and distribution; home ownership; prices of other goods and services in the area; social and cultural factors, such as schools and other institutions, and mass media; employment of the homemaker; and population density. In an analysis of the 1960-1961 Consumer Expenditure survey data, Lee and Phillips (1971) found differences in the consumption of many items between different urbanization groups.

In this study three different urbanization groups (urban, rural nonfarm and rural farm) are used to represent the package of urbanization characteristics that cannot be separated or measured. Each household was first classified as urban or rural using the U.S. Bureau of Census (1970, p. 82) guidelines where an urban household is defined as one located in an incorporated or unincorporated place of 2,500 or more inhabitants or in the densely settled urban fringe of an urbanized area. Within the rural category, a distinction between farm and nonfarm families was based on the presence of a farm operator.

Although there is no a priori basis for predicting the direction of effect of urbanization on toll expenditures, one would expect the difference in consumption between urban and rural nonfarm households to be greater than the consumption difference between rural farm and rural nonfarm households since the latter pair probably have more similar tastes.

Social Class. Behavioral scientists suggest that, due to a drive for conformity and the need for affiliation, an individual's behavior pattern may depend on his reference group, and that the main reference group as far as consumption patterns are concerned is one's social class. Education and occupation of the head of the household are used in this study as proxies for social class. 37

Education of the head was measured by years of formal schooling, including attendance at specialized business, trade, vocational and similar schools, completed at the time of the interview. To incorporate this variable into the analysis, the head of each household was placed into one of the following education categories: less than nine years of schooling, some high school, completed high school, some college, or college degree or beyond. It is possible that those individuals who have attended college may have more friends in distant cities than those achieving a lower educational level. Hence, the more educated may, ceteris paribus, do more long-distance calling.

The occupation of the head was also used as an indication of the social class of the family. The occupation of the family head was based on his major occupation, e.g., the occupation at which he was employed for the greatest number of weeks during 1974. The following occupational groupings were employed in this study: wage-salary workers, nonfarm self-employed, farm operators, and others not working. Ceteris paribus, the nonfarm self-employed and the farm operators may be expected to have higher long-distance bills due to placing business-related toll calls from their home telephone.

Race. Another factor which may affect a household's preferences and result in differences in consumption of toll message telephone service among households is race. Due to cultural differences, low asset holdings, or limited access to credit, blacks may exhibit expenditure patterns different from whites. Race has been found to be an important factor affecting the demand for other commodities. Telephone company executives suggest that blacks may spend more on toll calls than whites, everything else being the same.
Special Noneconomic Variables

All of the noneconomic variables considered thus far are standard taste variables of the type employed in most empirical demand studies. Some noneconomic variables that may specifically affect the demand for residential long-distance telephone service will now be considered.

The first three such variables are intended to reflect the hypothesis that a household does not place long-distance calls unless it has someone to call who is located outside its local calling area (i.e., the area within which telephone calls may be made without a toll charge). According to the Survey Research Center study (1956, p. 89, see footnote 15), most social long-distance calls are placed to either friends or relatives. The first of these variables is designed to measure the geographical dispersion of immediate family relatives living outside the local calling area and the other two variables are proxies for non-immediate relatives and friends located in distant cities. The fourth special noneconomic variable is intended to capture the idea that residential telephone subscribers who use the telephone intensively for local communication may also be high toll users. Each of these special noneconomic variables will now be discussed.

Geographical Dispersion of Family. At the conclusion of the interview each respondent was asked to list the relationships and the location of each member of the head and spouse's immediate family (e.g., parents, children and siblings) who lives outside the local calling area. This information was then used to construct two different specifications for geographical dispersion of family. One specification consisted of including the number of parents, number of children, and number of siblings living outside the local calling area as additional independent variables in the demand equation. An additional relative of each type, ceteris paribus, is expected to increase toll expenditures, although the magnitude of this effect may not be the same for each type of relative. The other specification for geographical dispersion of family, and the one used in the initial estimating equation, consists of entering into the demand equation the following three variables: total distance to all parents, total distance to all children, and total distance to all siblings. The distance-to-siblings variable, for example, represents the sum of the airline distances to the brothers and sisters that the head and spouse have living outside their local calling area. This variable thus takes into account the number of siblings as well as the distance to them. Intuitively, each of these three variables would be expected to have a positive effect on toll charges although the magnitude of this effect may not be the same for each type of relative.

Calling Scope. The second variable introduced into the demand function to take account of differences in "calling universe" or "potential callees" among the sampled households is calling scope, which is the number of main station telephones and private branch exchange trunks that a subscriber can call without paying a toll charge. An inverse relationship is hypothesized to exist between toll expenditures and calling scope, all other things remaining unchanged. As calling scope increases, a subscriber is able to call toll-free more friends and non-immediate relatives, business establishments, etc.

Recency of Residence. The third variable utilized in this study to adjust for differences in "potential callees" among households is the family's recency of residence. Those families that have recently moved into the area are more likely to have a larger number of friends and
relatives (other than immediate relatives) living outside their local
calling area than families that have lived in the same area for many
years. Thus, those families who have moved into the county after
January 1, 1970, would be expected to have higher long-distance
telephone bills than those families living in the same area for more
than five years.

Local Service Use. Residential telephone subscribers who make and
receive a large number of local telephone calls may have a preference
for communicating with friends and relatives by means of the telephone
rather than by visiting or writing letters. These people may also be
expected to be high toll users. Since no direct measure of local
service use is available, a proxy variable must be employed.

The household's total monthly local service charges, which was
part of the local service information provided by CT&T, was judged to
be a good indication of local service usage. The higher the quality of
grade of service the household receives and the more extension telephones
and other miscellaneous equipment that are rented, the higher the
monthly local service charges are, and the greater the probability
that the subscriber is a high local service user.\(^7\) A direct relation­
ship between total local service charges and toll expenditures would
exist if this hypothesis is correct.

Omitted Variables

In selecting variables for inclusion in a demand function, it is
not possible to include all sources of preference variation that may
influence consumption. For some factors of possible relevance data
are not available. In addition, it is desirable to make theories as
simple as possible and to include only variables of main importance.
Some of the factors that have been found to be important in studies of
the demand for other goods but which could not be included here in the
demand function for long-distance telephone calls are wealth, debt
status, home ownership status, income changes, buying intentions, and
confidence in the economy.

Since long-distance calling involves communicating with others,
psychological characteristics of the subscriber may be important
determinants of toll usage. As Currie (1972, p. 21) points out, "a
consumer purchases a commodity because he expects it to satisfy one or
more of his basic 'needs' or 'drives.' Some consumer needs are 'biogenic'
in origin and some 'psychogenic.'" Using data on 400 residential tele­
phone subscribers in southern Michigan, the Survey Research Center study
(1956, pp. 41-63, see footnote 15) analyzed the effects on toll expendi­
tures of three personality factors—security-insecurity, need for
affiliation, and underlying attitudes toward money. A multiple regres­
sion analysis revealed that none of these psychological variables exhibit
a significant effect on the probability of a household being a high user
of toll service. Due to the unavailability of the appropriate data,
similar variables could not be tested in this study.

Another possible source of variation in expenditures on toll calls
that could not be considered in this study is whether a family calls
their friends and relatives in distant cities or whether these people
call them. The Survey Research Center study (1956, pp. 26-27, see
footnote 15) found no tendency for either high or low toll users to be
recipients of many more long-distance calls than they make. Thus, the
omission of this factor is not considered to be important.

Another taste variable not included in the toll demand model is
whether the families studied have a preference for writing letters to
or visiting friends and relatives living away from home rather than
communicating by means of long-distance telephone calling. This
omission is not expected to be important. The Survey Research Center
study (1956, p. 25, see footnote 15) found that high toll users also write more letters and visit friends and relatives in distant cities more frequently than low toll users.

The influence of these and other factors that have not explicitly been included in the theoretical model gives rise to an error term in the estimating equation. As long as the omitted variables are not correlated with the included explanatory variables, the estimated parameters will remain unbiased.

Empirical Analysis

In the preceding part of this chapter a theoretical model was developed to explain household expenditures on long-distance calls. This demand function hypothesizes a household's expenditures on toll calls to be a function of the following variables: total gross family income, family size and composition, degree of urbanization, occupation and education of the family head, race, geographical dispersion of family, calling scope, recency of residence, and local service use. The purpose of this part of the chapter is to determine empirically if each of these explanatory variables has a significant effect on toll calling and to estimate the magnitude of these effects.

The outline of the empirical part of this chapter is as follows. The statistical technique selected for use and the data are discussed in the first two sections. The third section is devoted to determining the appropriate functional form for each of the quantitative explanatory variables. The empirical results are then presented and interpreted. Next, the extent to which the data fulfill the assumptions necessary to obtain the best linear unbiased parameter estimates is considered. The appropriateness of assuming the absence of interaction effects among the independent variables is analyzed in the concluding section of the chapter.

Selection of Statistical Technique

As in almost all econometric demand studies, the statistical technique here employed to estimate the demand function for toll calls is ordinary least squares regression analysis. There are several advantages of this technique. First, it provides results that indicate both the direction and magnitude of effect of each explanatory variable on the dependent variable. Secondly, it is a rather simple and inexpensive estimation technique. Thirdly, with a minimum of assumptions, ordinary least squares produces parameter estimates that are unbiased and that have the smallest variance in the family of linear unbiased estimators. The degree to which the data fulfill the assumptions necessary to obtain these results is examined later in this chapter.

Data

Data on a subset of the households interviewed in the Second Survey of Labor Supply in eastern North Carolina were used to estimate the demand function for residential long-distance telephone service. From mid-March to early June 1975, the Department of Economics and Business of North Carolina State University obtained completed interviews with a random stratified sample of 2,113 households living in eight eastern North Carolina counties. Of the 1,601 households living in single economic family housing units and receiving telephone service from Carolina Telephone and Telegraph Company, 71 percent gave written permission for the telephone company to release records on the telephone equipment installed in their homes, as well as detailed information on long-distance calls placed from their telephones. Of the 1,138 households for whom telephone records were obtained, 164 were deleted from the analysis for one or more of the following reasons: (1) the telephone installed in the home was classified as a
business main station, (2) one of the family members was a telephone company employee and thus entitled to concessions on both local and toll service, (3) someone else was paying the telephone bill, (4) toll charges were not available for each of the eight months included in the period of observation, or (5) data were missing for one or more of the explanatory variables. Thus, data on 934 households were available for use in estimating the toll demand function.

Since the aim of this analysis is not to obtain estimates of the mean or proportion for some variable for some aggregate of households, but rather to estimate the parameters of a demand equation, a random sample is not necessary. As Lansing and Morgan (1971, p. 233) point out, parameter estimates will remain unbiased even if the sample is non-random as long as there is no interaction between the explanatory variables and factors related to the probability of inclusion in the sample, i.e., as long as there are not factors associated with being included in the sample that alter the effect of an explanatory variable on the dependent variable.

Retaining in the data set only those households giving written permission for their telephone records to be released presents the possibility that there may be some factor associated with a household's signing of the permit slip that interacts with the explanatory variables altering their effect on expenditures on long-distance telephone calls. Because of this possibility the randomness or representativeness of the sample used in the toll demand analysis was analyzed. Since there are no independent outside sources of information with which the characteristics of the households included in the sample, or their expenditures on toll calls, can be compared, an alternative approach was employed. This procedure involved appraising the randomness of the sample at each stage in its selection process.

Since completed interviews were obtained with 85 percent of the occupied housing units listed in the stratified random sample of the eight-county area of eastern North Carolina, the initial sample of 2,113 households may be assumed to be a representative sample of the residential telephone subscribers living in the area surveyed. Next, to determine if those households used in the toll demand analysis are different from those households interviewed but not included in the analysis, the households in these two subgroups were compared with respect to a variety of characteristics, such as calling scope, geographical dispersion of family, income, urbanization, occupation, education, race, age, sex of head, etc. No major differences were found, suggesting that the sample used in the toll demand analysis is a random sample of the telephone subscribers in the area surveyed. Thus, it is not likely that the parameter estimates are biased due to an interaction between the explanatory variables and factors related to the probability of inclusion in the toll demand sample.

Selection of Functional Form

In order to estimate a demand equation it is first necessary to specify the mathematical form of the equation. Since ordinary least squares regression is the estimation technique employed, the relationship between the regressand and each explanatory variable, or some transformation of it, must be linear.

Income is the most important factor affecting a household's expenditure pattern. Thus, the first concern was to choose an appropriate mathematical form for the income-toll expenditures relationship. A separate regression equation was estimated for each of the seven functional forms discussed in the theoretical section on income (linear, quadratic, semi-log, hyperbolic, double-log, log-inverse, and log-log-inverse). Each of these equations included all
the other explanatory variables and reported income was used as the measure of purchasing power. Only data for white households were used in the estimations. On the basis of goodness of fit, significance of the individual coefficients, linearity of fit, and consistency with a priori notions, the double-log form was found to provide the best approximation to the true Engel curve.

A procedure suggested by Lansing and Morgan (1971, p. 309) was followed to determine the appropriate functional form for the relationship between toll charges and each of the following quantitative variables: total local service charges, calling scope, distance to parents, distance to children, and distance to siblings. Each of these continuous variables was converted into a set of dummy variables and entered into a multiple regression equation which included the other explanatory variables. An analysis of the coefficients of these dummy variables indicated that the double-log form provides a good approximation for the relationship between the dependent variable and each of these explanatory variables except calling scope. For this variable no simple relationship could be found to fit the data reasonably well. Calling scope was therefore grouped into classes and entered into the demand equation in the form of dummy variables. The following calling scope groups were used: less than 7,500; 7,500-10,000; 15,000-20,000; 20,000-25,000 and over 25,000. The effects of the qualitative explanatory variables, such as urbanization, education, occupation, race, and ancestry of residence, were taken into account by introducing into the regression equation dummy variables representing each category of each qualitative variable. To avoid complete multicollinearity, one category of each qualitative variable was deleted. Since the dependent variable in the estimating equation is the natural logarithm of toll expenditures, the antilogarithm of the estimated coefficient of a dummy variable indicates the average percentage difference in toll expenditures between a member of that dummy category group and a member of the left-out group similar in all other respects.55

As mentioned previously, including dummy variables in a single equation to take account of differences in tastes and preferences, rather than running separate regressions for each subgroup, implicitly assumes the absence of any interaction effects among the independent variables (i.e., that the effect of one independent variable on toll expenditures does not depend on the value of any other explanatory variable). The validity of this assumption is tested in the concluding section of this chapter.

In the preliminary regression equation the variables representing the number of family members in each age-sex category were entered in a linear form. The existence of zero values for some of these categories for many households precluded the use of a double-log form, and dummy variables for each category of each variable would greatly increase the independent variables. The antilogarithms of the estimated coefficients of these variables represent the percentage change in toll charges due to having an additional family member of a particular type, ceteris paribus. As Houthakker (1952, p. 16) points out, assuming the proportional effect of an additional family member in a particular age-sex category is constant implies that the absolute effect of an additional member will be greater the more members there are in that category.

One annoying drawback of working with a double-log functional form is that some method must be adopted for dealing with zero values of variables entered in the equation in logarithmic form. The presence of zero values for the dependent value was dealt with by adding $1 to total expenditures on toll calls for each observation before the
natural logarithm of expenditures was taken. For those households not having a parent, child, or sibling living outside the local calling area, the distance to that type of relative was set equal to one so that the logarithm of the distance would be equal to zero. 56

Now that the functional form for each of the continuous explanatory variables has been determined and the dummy variable categories for the qualitative variables have been defined, the estimating equation for expenditures on long-distance calls can be written as:

\[
\text{LTETERLD} = \alpha_0 + \alpha_1 \text{LTLSCSCH} + \alpha_2 \text{NIB} + \alpha_3 \text{N914} + \alpha_4 \text{N1517} + \alpha_5 \text{N1821} + \alpha_6 \text{N22} + \alpha_7 \text{F22} + \alpha_8 \text{SIMMENS} + \alpha_9 \text{COMP} + \alpha_{10} \text{SOMECOL} + \alpha_{11} \text{COMCPOL} + \alpha_{12} \text{FARMOP} + \alpha_{13} \text{UNEMPLOY} + \alpha_{14} \text{SELFEMPL} + \alpha_{15} \text{RURAL} + \alpha_{16} \text{LINC}74ES + \alpha_{17} \text{CS7.510} + \alpha_{18} \text{CS1520} + \alpha_{19} \text{CS2025} + \alpha_{20} \text{CS25} + \alpha_{21} \text{LDISTPAR} + \alpha_{22} \text{LDISTCHL} + \alpha_{23} \text{LDISTSBL} + \alpha_{24} \text{BLACK} + \alpha_{25} \text{MOVE},
\]

(15)

where the variables are defined in Table 2 and the \( \alpha \)'s are the parameters to be estimated.

Table 2. Definition of the variables used in the toll demand model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTETERLD</td>
<td>Natural logarithm of expenditures on long-distance calls during an eight-month period of 1975, in dollars.</td>
</tr>
<tr>
<td>LTLSCSCH</td>
<td>Natural logarithm of monthly total local service charges, in dollars.</td>
</tr>
<tr>
<td>N18</td>
<td>Number of household members 8 years of age or younger.</td>
</tr>
<tr>
<td>N914</td>
<td>Number of household members between the ages of 9 and 14, inclusive.</td>
</tr>
<tr>
<td>N1517</td>
<td>Number of household members between the ages of 15 and 17, inclusive.</td>
</tr>
<tr>
<td>N1821</td>
<td>Number of household members between the ages of 18 and 21, inclusive.</td>
</tr>
<tr>
<td>N22</td>
<td>Number of male household members 22 years of age or older.</td>
</tr>
<tr>
<td>F22</td>
<td>Number of female household members 22 years of age or older.</td>
</tr>
<tr>
<td>SOMEHS</td>
<td>1 if head completed more than 8 but less than 12 years of education; 0 otherwise.</td>
</tr>
<tr>
<td>COMPSH</td>
<td>1 if highest grade in school completed by head was the twelfth; 0 otherwise.</td>
</tr>
<tr>
<td>SOMECOL</td>
<td>1 if head completed 12 years of education or more; 0 otherwise.</td>
</tr>
<tr>
<td>COMP</td>
<td>1 if head completed 16 or more years of education; 0 otherwise.</td>
</tr>
<tr>
<td>FARMOP</td>
<td>1 if head is a farm operator; 0 otherwise.</td>
</tr>
<tr>
<td>SELFEMPL</td>
<td>1 if head is self-employed; 0 otherwise.</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>1 if head did not receive any employment income during 1974; 0 otherwise.</td>
</tr>
<tr>
<td>RURAL</td>
<td>1 if household is located in a rural area; 0 otherwise.</td>
</tr>
<tr>
<td>LINC74ES</td>
<td>Natural logarithm of estimated 1974 total gross family income, in dollars.</td>
</tr>
<tr>
<td>CS7.510</td>
<td>1 if 7,500 ≤ calling scope &lt; 10,000; 0 otherwise.</td>
</tr>
<tr>
<td>CS1520</td>
<td>1 if 15,000 ≤ calling scope &lt; 20,000; 0 otherwise.</td>
</tr>
<tr>
<td>CS2025</td>
<td>1 if 20,000 ≤ calling scope &lt; 25,000; 0 otherwise.</td>
</tr>
<tr>
<td>CS25</td>
<td>1 if calling scope ≥ 25,000; 0 otherwise.</td>
</tr>
<tr>
<td>LDISTPAR</td>
<td>Natural logarithm of the total distance to parents living outside the local calling area, in miles.</td>
</tr>
<tr>
<td>LDISTCHL</td>
<td>Natural logarithm of the total distance to children living outside the local calling area, in miles.</td>
</tr>
<tr>
<td>LDISTSBL</td>
<td>Natural logarithm of the total distance to siblings living outside the local calling area, in miles.</td>
</tr>
<tr>
<td>BLACK</td>
<td>1 if family is black; 0 if white.</td>
</tr>
<tr>
<td>MOVE</td>
<td>1 if family moved into county after January 1, 1970; 0 otherwise.</td>
</tr>
</tbody>
</table>

Regression Results

The results of estimating equation (15) using ordinary least squares regression are presented in Table 3, Model 1. To facilitate interpretation of these results, the sample means and standard deviations of the explanatory variables used in Model 1 are presented in Table 4. 57 The antilogarithm of the coefficient of each dummy and household composition
variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.

**Goodness of Fit.** The coefficient of determination ($R^2$) for Model 1, Table 3, indicates that 35 percent of the variation in the dependent variable is shown in Table 5. Table 6 contains a matrix of simple correlation coefficients between the explanatory variables.

The results given in Table 3, Model 1, are very encouraging. Only six of the 25 explanatory variables have coefficients that are not significantly different from zero at the 5 percent level using a two-tailed t-test. None of the variables with a significant coefficient has a sign opposite to that hypothesized in the theoretical sections.

Before the results for each of the explanatory variables are examined, the goodness of fit of Model 1 is evaluated.
### Table 4.
Sample means and standard deviations for the explanatory variables used in toll demand Model 1, Table 3

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTLSCH</td>
<td>2.20</td>
<td>.27</td>
</tr>
<tr>
<td>N18</td>
<td>.38</td>
<td>.76</td>
</tr>
<tr>
<td>N914</td>
<td>.35</td>
<td>.73</td>
</tr>
<tr>
<td>N1517</td>
<td>.23</td>
<td>.53</td>
</tr>
<tr>
<td>N1821</td>
<td>.30</td>
<td>.61</td>
</tr>
<tr>
<td>M22</td>
<td>.63</td>
<td>.49</td>
</tr>
<tr>
<td>P22</td>
<td>.28</td>
<td>.47</td>
</tr>
<tr>
<td>S60B8196</td>
<td>.37</td>
<td>.38</td>
</tr>
<tr>
<td>S60B8122</td>
<td>.26</td>
<td>.42</td>
</tr>
<tr>
<td>S66COL</td>
<td>.11</td>
<td>.31</td>
</tr>
<tr>
<td>S66COL</td>
<td>.11</td>
<td>.32</td>
</tr>
<tr>
<td>FARMOP</td>
<td>.09</td>
<td>.20</td>
</tr>
<tr>
<td>SELF'EMPL</td>
<td>.08</td>
<td>.27</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>.21</td>
<td>.41</td>
</tr>
<tr>
<td>RURAL</td>
<td>.64</td>
<td>.48</td>
</tr>
<tr>
<td>LINC74ES</td>
<td>9.14</td>
<td>.76</td>
</tr>
<tr>
<td>CS7.510</td>
<td>.23</td>
<td>.42</td>
</tr>
<tr>
<td>CS1520</td>
<td>.09</td>
<td>.28</td>
</tr>
<tr>
<td>CS2025</td>
<td>.25</td>
<td>.43</td>
</tr>
<tr>
<td>CS25</td>
<td>.17</td>
<td>.38</td>
</tr>
<tr>
<td>LDISTPAR</td>
<td>1.47</td>
<td>2.39</td>
</tr>
<tr>
<td>LDISTchs</td>
<td>2.59</td>
<td>2.99</td>
</tr>
<tr>
<td>LDISTHEL</td>
<td>4.92</td>
<td>2.70</td>
</tr>
<tr>
<td>BLACK</td>
<td>.28</td>
<td>.45</td>
</tr>
<tr>
<td>MOVE</td>
<td>.08</td>
<td>.27</td>
</tr>
</tbody>
</table>

### Table 5.
Antilogarithm of the coefficient of each dummy and household composition variable used in toll demand Model 1, Table 3

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Antilogarithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>N18</td>
<td>.987</td>
</tr>
<tr>
<td>N914</td>
<td>1.027</td>
</tr>
<tr>
<td>N1517</td>
<td>1.121</td>
</tr>
<tr>
<td>N1821</td>
<td>1.226</td>
</tr>
<tr>
<td>M22</td>
<td>1.363</td>
</tr>
<tr>
<td>P22</td>
<td>1.355</td>
</tr>
<tr>
<td>S60B8196</td>
<td>1.256</td>
</tr>
<tr>
<td>S60B8122</td>
<td>1.289</td>
</tr>
<tr>
<td>S66COL</td>
<td>1.587</td>
</tr>
<tr>
<td>S66COL</td>
<td>1.851</td>
</tr>
<tr>
<td>FARMOP</td>
<td>.915</td>
</tr>
<tr>
<td>SELF'EMPL</td>
<td>.986</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>.785</td>
</tr>
<tr>
<td>RURAL</td>
<td>.782</td>
</tr>
<tr>
<td>CS7.510</td>
<td>.704</td>
</tr>
<tr>
<td>CS1520</td>
<td>.638</td>
</tr>
<tr>
<td>CS2025</td>
<td>.560</td>
</tr>
<tr>
<td>CS25</td>
<td>.531</td>
</tr>
<tr>
<td>BLACK</td>
<td>1.513</td>
</tr>
<tr>
<td>MOVE</td>
<td>1.080</td>
</tr>
</tbody>
</table>
Table 6. Simple correlation coefficient matrix for the explanatory variables used in toll demand Model I, Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LnSCH</td>
<td></td>
<td>.163</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N18</td>
<td>.102</td>
<td></td>
<td>.090</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N914</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1517</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1821</td>
<td>.191</td>
<td></td>
<td>.091</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOMERS</td>
<td>.016</td>
<td>.005</td>
<td>.027</td>
<td>.082</td>
<td>.061</td>
<td>.086</td>
<td>.117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMPRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SONECOL</td>
<td>.075</td>
<td>.022</td>
<td>-.015</td>
<td>-.052</td>
<td>-.026</td>
<td>.000</td>
<td>-.089</td>
<td>.241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RURAL</td>
<td>.186</td>
<td>.119</td>
<td>.165</td>
<td>.131</td>
<td>.178</td>
<td>.645</td>
<td>.339</td>
<td>.078</td>
<td>.112</td>
<td>.532</td>
<td>.209</td>
<td>.020</td>
</tr>
<tr>
<td>CS1.510</td>
<td>-.078</td>
<td>.038</td>
<td>-.011</td>
<td>-.032</td>
<td>-.004</td>
<td>.008</td>
<td>.002</td>
<td>.015</td>
<td>-.035</td>
<td>-.050</td>
<td>-.036</td>
<td>.003</td>
</tr>
<tr>
<td>CS1.520</td>
<td>.058</td>
<td>.032</td>
<td>.010</td>
<td>.015</td>
<td>.036</td>
<td>.086</td>
<td>.049</td>
<td>.010</td>
<td>.033</td>
<td>-.010</td>
<td>-.016</td>
<td>.006</td>
</tr>
<tr>
<td>CS2.023</td>
<td>.089</td>
<td>.009</td>
<td>.030</td>
<td>.049</td>
<td>.055</td>
<td>.016</td>
<td>.010</td>
<td>.030</td>
<td>.007</td>
<td>.025</td>
<td>.013</td>
<td>-.058</td>
</tr>
<tr>
<td>CS2.035</td>
<td>.159</td>
<td>.012</td>
<td>.028</td>
<td>.055</td>
<td>.014</td>
<td>.013</td>
<td>.004</td>
<td>.030</td>
<td>.008</td>
<td>.006</td>
<td>.004</td>
<td>.001</td>
</tr>
<tr>
<td>LDISTPAR</td>
<td>.140</td>
<td>.117</td>
<td>.113</td>
<td>.029</td>
<td>.046</td>
<td>.145</td>
<td>.213</td>
<td>.088</td>
<td>.004</td>
<td>.117</td>
<td>.342</td>
<td>.067</td>
</tr>
<tr>
<td>LDISTCHL</td>
<td>-.020</td>
<td>-.261</td>
<td>-.104</td>
<td>.040</td>
<td>.085</td>
<td>-.072</td>
<td>.224</td>
<td>.058</td>
<td>-.180</td>
<td>-.112</td>
<td>-.115</td>
<td>-.044</td>
</tr>
<tr>
<td>LDISTBSL</td>
<td>.079</td>
<td>.063</td>
<td>.048</td>
<td>.100</td>
<td>.060</td>
<td>.130</td>
<td>.180</td>
<td>.002</td>
<td>.014</td>
<td>.110</td>
<td>.126</td>
<td>.047</td>
</tr>
<tr>
<td>BLACK</td>
<td>.126</td>
<td>.110</td>
<td>.153</td>
<td>.189</td>
<td>.163</td>
<td>-.078</td>
<td>.170</td>
<td>-.043</td>
<td>-.193</td>
<td>-.166</td>
<td>-.162</td>
<td>.001</td>
</tr>
<tr>
<td>MOVE</td>
<td>.065</td>
<td>.067</td>
<td>.027</td>
<td>-.023</td>
<td>.019</td>
<td>.000</td>
<td>-.121</td>
<td>-.048</td>
<td>-.023</td>
<td>.155</td>
<td>.195</td>
<td>-.064</td>
</tr>
</tbody>
</table>

Table 6 (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
<th>(16)</th>
<th>(17)</th>
<th>(18)</th>
<th>(19)</th>
<th>(20)</th>
<th>(21)</th>
<th>(22)</th>
<th>(23)</th>
<th>(24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LTLSCH</td>
<td>.165</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N914</td>
<td>.102</td>
<td>.090</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1517</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1821</td>
<td>.191</td>
<td></td>
<td>.091</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOMERS</td>
<td>.016</td>
<td>.005</td>
<td>.027</td>
<td>.082</td>
<td>.061</td>
<td>.086</td>
<td>.117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMPRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SONECOL</td>
<td>.075</td>
<td>.022</td>
<td>-.015</td>
<td>-.052</td>
<td>-.026</td>
<td>.000</td>
<td>-.089</td>
<td>.241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RURAL</td>
<td>.186</td>
<td>.119</td>
<td>.165</td>
<td>.131</td>
<td>.178</td>
<td>.645</td>
<td>.339</td>
<td>.078</td>
<td>.112</td>
<td>.532</td>
<td>.209</td>
<td>.020</td>
</tr>
<tr>
<td>CS1.510</td>
<td>-.078</td>
<td>.038</td>
<td>-.011</td>
<td>-.032</td>
<td>-.004</td>
<td>.008</td>
<td>.002</td>
<td>.015</td>
<td>-.035</td>
<td>-.050</td>
<td>-.036</td>
<td>.003</td>
</tr>
<tr>
<td>CS1.520</td>
<td>.058</td>
<td>.032</td>
<td>.010</td>
<td>.015</td>
<td>.036</td>
<td>.086</td>
<td>.049</td>
<td>.010</td>
<td>.033</td>
<td>-.010</td>
<td>-.016</td>
<td>.006</td>
</tr>
<tr>
<td>CS2.023</td>
<td>.089</td>
<td>.009</td>
<td>.030</td>
<td>.049</td>
<td>.055</td>
<td>.016</td>
<td>.010</td>
<td>.030</td>
<td>.007</td>
<td>.025</td>
<td>.013</td>
<td>-.058</td>
</tr>
<tr>
<td>CS2.035</td>
<td>.159</td>
<td>.012</td>
<td>.028</td>
<td>.055</td>
<td>.014</td>
<td>.013</td>
<td>.004</td>
<td>.030</td>
<td>.008</td>
<td>.006</td>
<td>.004</td>
<td>.001</td>
</tr>
<tr>
<td>LDISTPAR</td>
<td>.140</td>
<td>.117</td>
<td>.113</td>
<td>.029</td>
<td>.046</td>
<td>.145</td>
<td>.213</td>
<td>.088</td>
<td>.004</td>
<td>.117</td>
<td>.342</td>
<td>.067</td>
</tr>
<tr>
<td>LDISTCHL</td>
<td>-.020</td>
<td>-.261</td>
<td>-.104</td>
<td>.040</td>
<td>.085</td>
<td>-.072</td>
<td>.224</td>
<td>.058</td>
<td>-.180</td>
<td>-.112</td>
<td>-.115</td>
<td>-.044</td>
</tr>
<tr>
<td>LDISTBSL</td>
<td>.079</td>
<td>.063</td>
<td>.048</td>
<td>.100</td>
<td>.060</td>
<td>.130</td>
<td>.180</td>
<td>.002</td>
<td>.014</td>
<td>.110</td>
<td>.126</td>
<td>.047</td>
</tr>
<tr>
<td>BLACK</td>
<td>.126</td>
<td>.110</td>
<td>.153</td>
<td>.189</td>
<td>.163</td>
<td>-.078</td>
<td>.170</td>
<td>-.043</td>
<td>-.193</td>
<td>-.166</td>
<td>-.162</td>
<td>.001</td>
</tr>
<tr>
<td>MOVE</td>
<td>.065</td>
<td>.067</td>
<td>.027</td>
<td>-.023</td>
<td>.019</td>
<td>.000</td>
<td>-.121</td>
<td>-.048</td>
<td>-.023</td>
<td>.155</td>
<td>.195</td>
<td>-.064</td>
</tr>
</tbody>
</table>

Table 6 (Continued)
variable is explained by variation in the independent variables. While an $R^2$ of .35 is not particularly high, it is very satisfactory considering that the equation was estimated using ungrouped observations on individual households. The $F$ statistic for Model 1 reveals that the hypothesis that the coefficient of determination is equal to zero can be rejected at the one percent level of significance.

The inability of the regression equation to explain a large portion of the variation in the dependent variable is probably due to the many taste factors (particularly sociological, cultural, and psychological) that affect toll calling but are not included in Model 1. Mincer (1960, p. 470) and Reid (1960b, p. 474) believe that researchers should not be disturbed by low $R^2$'s obtained when individual families are studied, and that little purpose is served in trying to identify explanatory variables to account for all variation in expenditures. Since the goal of most demand studies is estimating the effect of consumption of income, household size and composition, occupation, education, urbanization, race, etc., many of the variables relating to the preferences and idiosyncrasies of individual families can be ignored as long as they are uncorrelated with the included explanatory variables.

The low $R^2$'s could also be the result of using measured current income instead of the theoretically relevant permanent income. Another explanation is that the eight-month period over which the households’ long-distance calling has been observed may not be representative of their typical behavior.

Income. Two different measures of total gross family income were available for use in the empirical estimations. Model 1, Table 3, presents the results using estimated income (LINC74ES), and Model 2 shows the estimates obtained using reported income (LINC74RP). Only those families revealing the interval into which their 1974 income fell could be used in estimating Model 2.

The estimated elasticity of expenditures on toll calls with respect to income (i.e., income elasticity), which is the coefficient of the income variable in a double-log demand function, is virtually identical for the two income measures (.18 for both). The $R^2$'s are also almost identical. Except for the coefficient of the dummy variable for self-employed heads (SELFEMP) and the coefficients of the education dummy variables, which are all higher in Model 1, the coefficients of the other explanatory variables are little affected by the income measure. Only estimated income will be used in additional estimations since this income measure is a continuous variable and because it permits a larger size sample to be used. In addition, poor fits were achieved in the lower income ranges using reported income.

Income elasticities derived from cross-sectional household budget data are usually interpreted to be estimates of long-run income elasticities. In other words, they measure how a “typical consumer” will respond to a change in income assuming the consumer has time to adjust his consumption patterns fully to the new level of income. Since toll expenditures by some families in the sample may not be in equilibrium with respect to their current measured income, it may not be appropriate to interpret this coefficient as the long-run income elasticity. However, Currie et al. (1972, p. 133) point out that for normal patterns of income changes, it is more likely to be nearer the long-run parameter than the short-run parameter.

The magnitude of the estimated income elasticity of demand for toll calls in Models 1 and 2 is surprisingly low. Most services have income elasticities greater than one. If one were to classify commodities as
necessities or luxuries on the basis of an income elasticity less than or greater than unity, an income elasticity of .19 would imply that long-distance telephone calling is a necessity. A priori, one would not expect toll calls to have an income elasticity lower than those usually found for food and housing.

Although there are no cross-sectional estimates of the income elasticity of demand for residential toll calls with which to compare the estimates found here, a few income elasticities have been calculated using time-series data. Dobell et al. (1972) applied the stock-adjustment model of Bouthakker and Taylor (1970, pp. 9-26) to time-series data for Bell Canada and estimated the short-run and long-run income elasticities of demand for residential toll calls to be .20 and 1.3, respectively. Waverman (1973, see footnote 9), believing the demand function of Dobell et al. to be misspecified, reformulated the model and estimated it using the same data set. Waverman calculated the short-run income elasticity to be 1.11. Thus, the income elasticity calculated here is substantially lower than the best available estimate.

One possible explanation for the low income elasticity of demand for toll calls obtained in this study is that current measured income, rather than permanent income, is used as the measure of purchasing power. As mentioned previously, when current income is used, the calculated income elasticity may underestimate the theoretically appropriate permanent income elasticity. The extent of this downward bias is not believed to be very great. Carliner (1973) found that using measured income instead of permanent income produced underestimates of the income elasticity of demand for housing of less than .13. Lee (1968) estimated the permanent income elasticity of demand for housing to be .24-.37 higher than the estimates of the measured income elasticity of demand. The difference between the current and permanent income elasticity for a service such as long-distance telephone calls would not be expected to be this great. In an analysis of Canadian budget data, Ashikapoulos (1965) could not find a significant difference between the current and permanent income elasticity for any commodity group studied. Thus, it cannot be concluded that the low income elasticity of demand for long-distance telephone service obtained in this study is due solely to using current rather than permanent income in the calculations. The implications of such a low income elasticity for regulation and planning by telephone companies are discussed in the concluding chapter of this dissertation.

Occupation. Turning now to the set of dummy variables representing the occupation of the head of the household, it can be seen from Models 1 and 2 that families in which the head is a farm operator or is self-employed do not, contrary to a priori expectations, spend a significantly different (at the 10 percent level) amount on toll calls than families headed by a wage-salary worker. The results do show that families with a head receiving no employment income during 1914 spent 21 percent more on toll calls during 1915 than wage-salary headed families similar in all other respects.

Recency of Residence. The empirical results presented in Table 3, Models 1 and 2, also reject (at the 10 percent level of significance) the hypothesis that families moving into the county after January 1, 1970, spend more on long-distance calls than families living in the same county for five or more years, ceteris paribus. If this variable can be interpreted as a proxy for the number of friends and nonimmediate relatives a household has living outside its local calling area, the insignificance of MOVE and the significant coefficients on the distance-to-relatives variable can be regarded as a verification of the finding by the Survey Research Center study (1956, p. 23, see footnote 15) that
"Social calls to friends are less important than calls to relatives." In that study residential telephone subscribers in southern Michigan revealed that they were more likely to write to keep in touch with friends but to call to keep in touch with relatives living in distant cities.

Household Size and Composition. It is clear from Models 1 and 2, Table 3, that family size and composition are important determinants of long-distance calling. The coefficients of N1821, M22, and F22 are highly significant and have the expected signs. The coefficients of the variables representing the younger age groups (i.e., N18, N914, and N1517) are not significantly different from zero at the 5 percent level, indicating that additional family members in these age categories have no effect on toll expenditures. This is not surprising and agrees with the a priori hypothesis that the specific effect (i.e., the need for toll telephone service) for adolescent family members is less than that for adults.

Model 1 results suggest that an additional family member in the N1821, M22, and F22 age-sex groups may not have the same effect on toll expenditures. The appropriate procedure for testing the null hypothesis that the coefficients of these three variables are equal is the F-test. Since this test statistic is employed many times in this study, the procedure followed in using the F-test to test the null hypothesis that the coefficients of N1821, M22, and N22 are equal will be described in detail.

The first step in this test involves estimating the following two regression equations: (1) a restricted regression which incorporates the null hypothesis that the coefficients of N1821, M22, and F22 are equal; and (2) an unrestricted regression which imposes no restrictions on the coefficients of these three household composition variables. The results of estimating the unrestricted regression equation are given in Model 3, Table 3. The three nonsignificant family composition variables (i.e., N18, N914, N1517) and the three other insignificant variables (i.e., PARMOP, SELFEMPL, MOVES) contained in Model 1, Table 3, were omitted from the list of independent variables contained in Model 3.64 Model 4, Table 3, shows the results of estimating the restricted regression equation representing the null hypothesis. In this equation, PPH18 is the number of household members 18 years of age or older (i.e., PPH18 = N1821 + M22 + F22).

The second stage in this testing procedure involves calculating an F statistic, given below as:

\[
F = \frac{\text{SSE} - \text{SSE}}{r} \times \frac{\text{SSE}}{T - K - 1}
\]

where:

- \( \text{SSE} \) is the sum of squared residuals for the regression incorporating the null hypothesis (i.e., the restricted regression),
- \( \text{SSE} \) is the sum of squared residuals for the regression imposing no restrictions on the coefficients (i.e., the unrestricted regression),
- \( r \) is the number of independent restrictions imposed by the null hypothesis, and
- \( T - K \) is the degrees of freedom for the unrestricted regression.

From Table 3 it can be seen that the residual sum of squares for the unrestricted regression (\( \text{SSE} \)) is 863.3525 and that the residual sum of squares for the restricted regression equation (\( \text{SSE} \)) is 863.5967. The number of independent restrictions (\( r \)) imposed in Model 4 is 2 and the degrees of freedom (\( T - K \)) corresponding to Model 3 are 914, so that:

\[
F = \frac{2.2442/2}{863.3525/914} = 1.19
\]
Under the null hypothesis that $N_{21} = M_{22} = F_{22}$, $F$ is distributed as an $F$ distribution with $r$ degrees of freedom in the numerator and $T - K$ degrees of freedom in the denominator. Thus, the test criterion is to reject the null hypothesis at a desired level of significance (usually 1 percent or 5 percent) if $F > F_c$, where $F_c$, the critical value, is the value of the $F$ distribution for the specified level of significance. Corresponding to two degrees of freedom in the numerator and 914 degrees of freedom in the denominator, $F_c$ equals 3 for the 95 percent level of confidence. Since $F < F_c$, the null hypothesis that the coefficients of $N_{21}$, $M_{22}$, and $F_{22}$ are equal cannot be rejected at the 5 percent level of significance.

A re-estimation of Model 4, Table 3, using dummy variables for families with 2, 3, and 4 members 18 years of age or older indicated that the relationship between the logarithm of toll charges and $PPH_{18}$ is linear. Thus, Model 4, Table 3, incorporates the appropriate functional form for $PPH_{18}$ and implies that each additional household member 18 years of age or older, $ceteris paribus$, increases expenditures on long-distance calls by 25 percent.

**Education.** The coefficients of the dummy variables representing the education of the head in Table 3 are all highly significant. The magnitude of these coefficients suggests that toll expenditures increase as education of the head increases. However, F-tests, similar to the one described above, indicate that there is not a significant (at the 5 percent level) difference between the coefficients of $SOHEHS$ and $COMPHS$ or between the coefficients of $SOMECOL$ and $COMPCOL$. There is a significant (at the 1 percent level) difference in toll expenditures between families with a head completing between nine and 12 years of school, inclusive, and households with a head completing more than 12 years of education. A re-estimation of Model 4, which is not presented here, reveals that these two groups spend 29 and 73 percent more on toll calls, respectively, than households headed by individuals completing eight years or less of schooling.

The direct relationship between toll expenditures and the education of the head may be the result of families with a college-educated head having more friends and nonimmediate relatives living outside their local calling area. Alternatively, education may be serving as a proxy for permanent income. This finding may also simply reflect differences between these groups in tastes and preferences for long-distance telephone calling.

**Race.** The coefficients of the dummy variable for race (BLACK) in Table 3 are significantly different from zero at the 1 percent level and verify the telephone company executives' hypothesis that blacks spend more on long-distance calling than whites. Model 4, Table 3, indicates that black families spend a surprising 56 percent more on toll calls than white families in similar circumstances. This may reflect a preference by blacks for long-distance calling versus other means of communication, and indicates that the integrity of black families is greater than commonly believed.

**Calling Scope.** Looking now at the dummy variables representing the calling scope groups, Table 3 shows that the coefficient of each of these variables has the expected sign and is significantly different from zero at the 1 percent level. The magnitude of these coefficients suggests that as calling scope increases, $ceteris paribus$, expenditures on toll calls decrease initially, then increase, and then decrease again. However, an F-test could not reject (at the 5 percent level of significance) the null hypothesis that the coefficients of $CS1520$, $CS2520$, $CS3520$, $CS4520$, and $CS5520$ are all equal.
CS2025 and CS25 are equal. Thus, it cannot be concluded that the empirical results are inconsistent with the a priori expectation of an inverse relationship between toll expenditures and calling scope.

Urbanization. The coefficient of the urbanization dummy variable (RURAL) is significantly different from zero at the 1 percent level in all equations contained in Table 3. Model 3 results indicate that rural households spend 22 percent less on toll calls than similar families living in urban areas.

Dummy variables for the two rural subdivisions, rural farm and rural nonfarm, could not be entered into the same equation with FARMOP since the definition of a farm operator and a rural farm are so similar. When Model 4, which does not contain FARMOP, was re-estimated with RURAL deleted and dummy variables for rural farm and rural nonfarm included, the coefficients of both these variables were found to be negative and significantly different from zero at the 1 percent level. However, the null hypothesis that the coefficients of these two variables are equal could not be rejected (at the 5 percent level of significance) using an F-test. Thus, the demand functions estimated in Table 3 incorporate the appropriate specification for urbanization.

Local Service Use. The regression results presented in Table 3 provide empirical support for the hypothesis that subscribers with higher monthly bills for local service also have higher toll charges. In all equations the coefficient of LTLSC is positive and highly significant. If total local service charges can be regarded as a proxy for the amount that the telephone is used for local communication, these results suggest that high local service users are also high toll users.

An alternative measure of local service use was also examined. With this specification dummy variables were used to represent the grade of service the subscriber is receiving and whether or not the family is renting an extension telephone. The hypothesis is that the more intensively the telephone is used for transmitting and receiving local messages, the greater is the probability that the subscriber will have a higher grade of residential main station telephone service and the more likely the subscriber is to rent an extension telephone. To test this alternative specification, Model 4, Table 3, was re-estimated after deleting LTLSC and adding the following three dummy variables:

- RMT2 = 1 if the household rents a two-party residential main station; 0 otherwise.
- RMT4 = 1 if the household rents a four-party residential main station; 0 otherwise, and
- RET = 1 if the household rents one or more extension telephones; 0 otherwise.

The results of estimating this respecified equation are judged to be of interest and are presented in Table 7. The estimated coefficients of the three new proxy variables for local service use have the anticipated sign and are significantly different from zero at the 5 percent level. Although the magnitude of the coefficients of RMT2 and RMT4 suggest that subscribers with four-party lines spend less on toll calls than those with two-party lines, an F-test failed to reveal a significant difference (at the 5 percent level) between these coefficients. Regression results not shown here indicate that households with a two- or four-party main station spend, on the average, 30 percent less on toll calls than families with a private line, ceteris paribus. Table 7 suggests that residential telephone subscribers having one or more extension telephones spend 28 percent more on long-distance calls than similar families without an extension station.

Geographical Dispersion of Family. In Table 3 and Table 7 the coefficients of the three variables representing the natural logarithm
Table 7. Regression results for the toll demand model using an alternative specification for local service use

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.1998*</td>
<td>.5029</td>
</tr>
<tr>
<td>PPHL</td>
<td>.2281**</td>
<td>.0445</td>
</tr>
<tr>
<td>SOMEHS</td>
<td>.2365*</td>
<td>.0944</td>
</tr>
<tr>
<td>CONFHS</td>
<td>.3326*</td>
<td>.0926</td>
</tr>
<tr>
<td>SOMECOIL</td>
<td>.3774*</td>
<td>.1235</td>
</tr>
<tr>
<td>CONFCOIL</td>
<td>.5482**</td>
<td>.1304</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>-.2814**</td>
<td>.0917</td>
</tr>
<tr>
<td>RURAL</td>
<td>-.0149</td>
<td>.0826</td>
</tr>
<tr>
<td>LINC74+</td>
<td>.1587**</td>
<td>.0571</td>
</tr>
<tr>
<td>CS7.510</td>
<td>-.2930**</td>
<td>.0926</td>
</tr>
<tr>
<td>CS1520</td>
<td>-.7166**</td>
<td>.1250</td>
</tr>
<tr>
<td>CS2025</td>
<td>-.5125*</td>
<td>.0954</td>
</tr>
<tr>
<td>CS25</td>
<td>-.4453**</td>
<td>.1099</td>
</tr>
<tr>
<td>LDISTPAR</td>
<td>.0839**</td>
<td>.0133</td>
</tr>
<tr>
<td>LDISTCHL</td>
<td>.1230**</td>
<td>.0118</td>
</tr>
<tr>
<td>LDISTSBL</td>
<td>.0411**</td>
<td>.0125</td>
</tr>
<tr>
<td>BLACK</td>
<td>.4256**</td>
<td>.0829</td>
</tr>
<tr>
<td>RET</td>
<td>.5433**</td>
<td>.0703</td>
</tr>
<tr>
<td>RMT2</td>
<td>-.2571*</td>
<td>.1081</td>
</tr>
<tr>
<td>RMT4</td>
<td>-.4051**</td>
<td>.0928</td>
</tr>
</tbody>
</table>

| R²                   | .3578       |
| F                    | 26.7989     |
| SSEb                 | 851.6361    |

*Indicates significance at the 5% level and **indicates significance at the 1% level using a two-tailed t-test.

<table>
<thead>
<tr>
<th>NOPAR1</th>
<th>1 if one parent lives outside the local calling area; 0 otherwise,</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOPAR2</td>
<td>1 if two or more parents live outside the local calling area; 0 otherwise,</td>
</tr>
<tr>
<td>NOCHILL</td>
<td>1 if one child lives outside the local calling area; 0 otherwise,</td>
</tr>
<tr>
<td>NOCHILL2</td>
<td>1 if two or more children live outside the local calling area; 0 otherwise,</td>
</tr>
</tbody>
</table>

The magnitude of the coefficients of LDISTPAR, LDISTCHL, and LDISTSBL reveals that a family will increase its long-distance calling the most as the distance to its children increases and the least as the distance to siblings increases, everything else remaining the same. An F-test rejects at the 5 percent, but not the one percent, level of significance the null hypothesis that the coefficients of LDISTPAR and LDISTCHL are equal. The hypothesis that the coefficients of all three variables are equal can be rejected at the one percent level.

An alternative measure of the geographical dispersion of family, discussed in the theoretical section of this chapter, is the number of parents, children, and siblings living outside the local calling area. No simple functional form for the relationship between each of these three variables and toll expenditures could be found. Thus, dummy variables were used to introduce these variables into the demand equation. A series of F-tests indicated that the appropriate categories for the dummy variables are the following:
NOSIBL2 = 1 if two or more siblings live outside the local calling area; 0 otherwise.

The results of re-estimating Model 4, Table 3, with the three distance-to-relatives variables deleted and the above dummy variables added are presented in Table 8. The coefficients of all these variables are highly significant and have the anticipated sign. The relative magnitude of these coefficients is consistent with the earlier findings using the distance-to-relatives specification.

Several attempts to determine the effects on toll expenditures of an increase in the total distance to a particular type of relative, the number of relatives held constant or, alternatively, the effects of more relatives of a particular type, the total distance to them remaining the same, were not successful. Including the number of each type of relative in an equation containing the total distance to each type of relative or in an equation with the average distance to each type of relative resulted in several of the variables becoming nonsignificant due to severe multicollinearity. None of the variables, however, have negative coefficients, indicating that toll expenditures, if anything, increase with either an increase in the number of relatives, distance to relatives remaining the same, or an increase in distance to relatives, the number of relatives remaining constant.

Appropriateness of the Statistical Technique

The analytical technique employed in the empirical analysis of the demand for residential long-distance telephone service is a very powerful and widely used statistical tool. The power of this method depends on the underlying assumptions being fulfilled in the particular application. Ordinary least squares regression analysis will provide best (i.e., minimum variance) linear unbiased estimators assuming the error terms have a mean of zero, the same variance, and are distributed independently of each other and the explanatory variables. The extent to which these assumptions are fulfilled with the data used in the toll demand analysis will now be examined.

Table 8. Regression results for the toll demand model using an alternative specification for geographical dispersion of family

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.028</td>
<td>.5462</td>
</tr>
<tr>
<td>LTLSCH</td>
<td>.6651**</td>
<td>.1298</td>
</tr>
<tr>
<td>PPHLA</td>
<td>.1211**</td>
<td>.0447</td>
</tr>
<tr>
<td>SOMEKES</td>
<td>.2588**</td>
<td>.0962</td>
</tr>
<tr>
<td>COMPHS</td>
<td>.2659**</td>
<td>.0928</td>
</tr>
<tr>
<td>SONECOFL</td>
<td>.5135**</td>
<td>.1213</td>
</tr>
<tr>
<td>COMPCOL</td>
<td>.6603**</td>
<td>.1279</td>
</tr>
<tr>
<td>UNEMPLOYE</td>
<td>-1.2591**</td>
<td>.0919</td>
</tr>
<tr>
<td>RURAL</td>
<td>-1.2727**</td>
<td>.0756</td>
</tr>
<tr>
<td>LINCOME</td>
<td>.1680**</td>
<td>.0569</td>
</tr>
<tr>
<td>CS7.510</td>
<td>-.3055**</td>
<td>.0528</td>
</tr>
<tr>
<td>CS1520</td>
<td>-.7612**</td>
<td>.1261</td>
</tr>
<tr>
<td>CS2025</td>
<td>-.3547**</td>
<td>.0970</td>
</tr>
<tr>
<td>CS25</td>
<td>-.5834**</td>
<td>.1133</td>
</tr>
<tr>
<td>BLACK</td>
<td>.3192**</td>
<td>.0827</td>
</tr>
<tr>
<td>UNPA1X</td>
<td>.5972**</td>
<td>.0990</td>
</tr>
<tr>
<td>UNPA12</td>
<td>.6095**</td>
<td>.0992</td>
</tr>
<tr>
<td>UNCHIL1</td>
<td>.5114**</td>
<td>.0883</td>
</tr>
<tr>
<td>UNCHIL2</td>
<td>.8961**</td>
<td>.0820</td>
</tr>
<tr>
<td>NOSIBL2</td>
<td>.2666**</td>
<td>.0711</td>
</tr>
</tbody>
</table>

\[ R^2 = .3573 \]
\[ F = 26.7465 \]
\[ SSE = 852.4320 \]

a* indicates significance at the 5% level and ** indicates significance at the 1% level using a two-tailed t-test.
bResidual sum of squares.
The assumption from the above list that is most commonly violated in cross-sectional demand studies is constant variance of the error terms. If heteroscedasticity is present, the parameter estimates will remain unbiased although they will no longer be minimum variance. The variance of the error term may increase, as suggested by Houthakker (1952, p. 3) and Kmenta (1971, p. 249), as income increases if, at higher income levels, tastes have a more important effect on expenditures than income and the other explanatory variables. A plot of the residuals from Model 1, Table 3, against LINC74ES revealed no tendency for the variance of the residuals to be related to the natural logarithm of estimated income.

A more formal test for the existence of heteroscedasticity, appropriate for this large sample size, was also conducted. Following Goldberger (1964, p. 246), the sample variance of the residuals was calculated for arbitrarily marked off ranges of LINC74ES. A chi-square test could not reject (at the 5 percent level of significance) the null hypothesis that these variances are equal. It may be, as suggested by Hamborg (1960, p. 353), that any lack of consistency between the variance of expenditures and income has been reduced by taking the natural logarithm of expenditures and income.

Although the assumption that the error terms are independently distributed random variables is more frequently violated when time-series rather than cross-sectional data are being analyzed, conceptually the problem exists for both types of data. The clustered nature of the sample utilized in this study, resulting from interviewing samples of housing units within small areas of land referred to as segments, may produce intercorrelation among the error terms. The presence of such serial correlation does not affect the unbiasedness or consistency of the estimators, although it does make them less efficient. The standard errors of the parameter estimates presented in Tables 3, 7, and 8 assume that the households interviewed were selected independently. That is, they assume a simple random sample. As a consequence of clustering, the number of independent observations is not the same as the number of households. Thus, the reported standard errors are underestimates of the actual or true standard errors. Some of the explanatory variables with coefficients reported to be significantly different from zero may not actually be significant if the true standard errors were used to calculate the t-statistics.

There is no direct way to calculate standard errors for regression coefficients from complex samples. However, Kish and Frankel (1970) describe and justify what they call a balanced repeated replication method for computing standard errors of regression coefficients. When this complicated procedure was used to calculate standard errors for regression coefficients in five empirical studies, the ratio of actual standard errors to simple random sampling standard errors ranged from 1.05 for less clustered to 1.46 for more clustered samples. Kish and Frankel conclude that standard errors computed by machine programs, based on simple random sampling assumptions, were not gross underestimates. The calculation of standard errors using the balanced repeated replication procedure was beyond the scope of the present study. The effects of clustering are not considered to be of primary importance in this study and are not allowed for.

Ordinary least squares also assumes the absence of any linear dependence between the explanatory variables. If some of the explanatory variables are highly, but not perfectly, collinear, it is difficult to distinguish the effect of each of these variables on the dependent variable with any real precision and the standard errors of the coefficients will therefore be large. Although the presence of a high
degree of multicollinearity does not imply that the correlation between
any two explanatory variables must be particularly high, the simple
correlation coefficient matrix in Table 6 suggests that multicollinearity
is not a problem. This is not unusual for samples of this size.

Test for Interactions

The long-distance demand models developed and estimated in this
chapter assume that the effect of each explanatory variable on toll
charges does not depend on the level or value of other independent
variables (e.g., it is assumed that there are no interactions among
the explanatory variables). Such interactions have been found to
classify the demand for other commodities. Since the qualitative
explanatory variables were found to exert a significant effect on the
intercept of the demand equation, the effect of these variables on the
coefficients of the other variables was examined.

To investigate the possibility of interactions, a separate regres­sion
equation was estimated for each category of each qualitative
variable. All the other explanatory variables were included in each of
these equations. Thus, a separate demand equation was estimated for
blacks, whites, urban households, rural households, families in each
of the education, occupation, and calling scope categories, and for
households with 1, 2, 3, and 4 persons 18 years of age or older.

For each qualitative variable an F-test was performed to test the
null hypothesis that all the parameters, except the intercept, are
equal in the regressions for each category of the variable. This
hypothesis could be rejected (at the 5 percent level of significance)
only for the urbanization variable. Thus, there is almost complete
absence of interactions among the independent variables, and, except
for urbanization, the effects of the qualitative variables are
restricted to the intercept.

4

Demand for Residential
Extension Stations

In this chapter the demand for residential extension stations is
studied. Demand theory is used first to develop briefly a
theoretical model of the factors hypothesized to affect a residential
subscriber's decision to rent an extension telephone. The development
of this model is followed by a discussion of alternative statistical
techniques appropriate for the situation investigated. In the final
section the results of estimating the model using binary logit analysis
and ordinary least squares regression are presented and interpreted.

Data on a subset of the Carolina Telephone and Telegraph Company resi­
dential subscribers interviewed in the Second Survey of Labor Supply in
Eastern North Carolina were examined in the empirical analysis.
Theoretical Model

As in the toll demand analysis, the appropriate unit of observation for studying the demand for residential extension telephones is the household. The phenomenon to be explained by the theoretical model is whether or not a residential telephone subscriber chooses to rent one or more inside extension telephones. In this analysis there is what is usually referred to as a dichotomous, binary, quantal, or (0,1) dependent variable.

No matter how many explanatory variables are included in the model, there will still be many others that may influence choice in an individual case. A household’s decision cannot be predicted with complete certainty. Thus, the theoretical model developed here is best interpreted as determining the factors affecting the probability that a residential subscriber will rent one or more extension telephones. The independent variables considered may be grouped into three categories—economic variables, traditional noneconomic variables, and noneconomic variables specific to the service under analysis.

Economic Variables

Classical utility theory suggests that income and prices should be important determinants affecting a household’s decision to rent an extension station. The measure of purchasing power employed in this analysis is estimated total family gross income in 1974. The same measurement and theoretical difficulties encountered in using current measured income in the toll demand analysis are applicable to this problem. These difficulties will not be repeated here. A priori, an increase in income, everything else being the same, is expected to increase the probability of a household having an extension telephone.

Since the data used in the empirical analysis are cross-sectional, the price of an extension telephone cannot be treated as an additional independent variable. All households pay the same $1.25 monthly rental rate for a company-provided extension telephone. Also, at the time this study was conducted a residential telephone customer could not legally install a noncompany extension telephone without the purchase of a protective device from the telephone company. Thus, substitutes are assumed not to exist.

Traditional Noneconomic Variables

In order to understand consumer behavior, as well as to obtain unbiased estimates of the effects of income changes, traditional noneconomic variables representing tastes and preferences are introduced into the model as additional independent variables. The variables included in this group are the following: household size and composition, degree of urbanization, and race, education, and occupation of the head of the household. The same specification and categories for these variables that were used in the toll demand model of Chapter 3 are again employed here. The expected direction and magnitude of effect of these variables on the probability of renting an extension telephone will now be considered.

An additional family member in each age-sex group should, ceteris paribus, increase the household’s probability of renting an extension telephone, assuming the specific effect of an additional member outweighs the income effect. The total effect of younger family members is expected to be smaller than the total effect of older members since young people add less to total household need for extension telephones (i.e., have a smaller specific effect).

A priori, one cannot predict the direction of effect that dummy variables for education, occupation, or urbanization will have.
Discussions with telephone company executives suggest that blacks, ceteris paribus, may be more likely to have extension telephones than whites.

Special Non-economic Variables

In addition to the above standard independent variables of the type used in most demand studies to take account of differences in tastes and preferences, two special non-economic factors are hypothesized to affect the demand for the particular service under analysis. The first of these variables, size of housing unit, is introduced into the model to capture the idea that the larger the housing unit, the greater the utility a family would receive from an extension telephone, and thus the more likely it would be to rent one, ceteris paribus. A proxy variable, number of levels of living quarters, is used to measure the size of the housing unit.

The second special non-economic variable hypothesized to affect the probability of a household renting an extension telephone is the intensity with which the telephone is used for making and receiving local calls. The proxy variable used to measure local service usage is net local service charges, calculated by subtracting monthly charges for renting extension telephones from total local service charges. Total local service charges were defined and used in Chapter 3. The hypothesis is that the more local calls a family makes and receives, the higher its net local service charges may reasonably be expected to be, the greater utility it would receive from an extension telephone, and ceteris paribus, the more likely it would be to rent one.

Data

Data on a subset of the households interviewed were used to test the theoretical model developed in the previous section. Of the 1,601 CT&T subscribers interviewed who lived in a single economic family housing unit, 71 percent gave written permission for the telephone company to release their telephone records. Of these 1,138 households, 120 were deleted from the data set for one of the following reasons: (1) none of the local service information was obtained, (2) the telephone installed in the home was classified as a business main station, (3) a member of the household was a telephone company employee, (4) someone else was paying the telephone bill, or (5) data were missing for one or more of the independent variables. Of the 1,018 remaining households, 42 percent were renting one or more inside extension telephones on June 20, 1975, the day on which the local service information was obtained from the telephone company.

Although randomness of the sample is not essential for estimating relationships, as mentioned in the previous chapter, biased estimates of the effect of some explanatory variables may be obtained if there are interactions between the independent variables and variables related to the probability of inclusion in the sample. Such an interaction could result from using data only for those households signing the permit slip. Also, with the statistical technique used to estimate the extension demand function, a random sample is useful for forecasting, as will be discussed later. For these two reasons, the representativeness or randomness of the sample was evaluated using two different approaches. The first approach used to measure the representativeness of the sample involved comparing information available from the survey with information from an independent outside source. The number of residential extension stations per household for the sample used in this analysis was compared with the number of residential extension telephones per residential main station telephone, calculated for CT&T exchanges located within the eight-county survey area, and was found to be within...
a 95 percent confidence interval of the latter figure. The second method used to assess the randomness of the sample involved comparing the general characteristics of the households included in the extension demand sample with the characteristics of the residential telephone subscribers interviewed but not included in the data set. This comparison revealed few differences between the two groups. The high response rate (85 percent) obtained in the survey, which was designed using a stratified random sampling technique, together with the finding, provides additional support for assuming the sample to be random. Thus, the parameter estimates are not expected to be biased due to an interaction between the independent variables and variables related to the probability of inclusion in the sample.

Selection of Statistical Technique

The data set described in the previous section provides, for each observation, values for each of the specified explanatory variables and knowledge of whether or not the household rents one or more extension telephones. The statistical problem is to estimate the relationship between the explanatory variables and the choice that is made. This situation is similar to estimating the demand for houses or automobiles, in the sense that the consumer is choosing among discrete alternatives, rather than different quantities of divisible goods. This results in what is referred to by Warner (1962) as a binary choice problem.

The theoretical model of the demand for residential extension telephones developed in the first part of this chapter hypothesizes that the probability of a residential telephone customer renting one or more extensions is a function of household income, household size and composition, degree of urbanization, race, education and occupation of the household head, size of housing unit, and local service usage. Letting \( \text{RET} = 1 \) if a household rents an extension telephone and letting \( \text{RET} = 0 \) if it does not, the theoretical model can be expressed in functional notation as follows:

\[
\Pr (\text{RET} = 1) = \phi (X' \beta)
\]

where \( X \) is a column vector of values for the explanatory variables for the household and \( \beta \) is a column vector of parameters to be estimated such that

\[
X' \beta = \beta_0 + \beta_1 \text{NIL8} + \beta_2 \text{NIL14} + \beta_3 \text{NIL157} + \beta_4 \text{NIL821} + \beta_5 \text{M22} + \beta_6 \text{F22} + \beta_7 \text{SOMERS} + \beta_8 \text{COMPUS} + \beta_9 \text{SOMECOL} + \beta_{10} \text{COMPUS} + \beta_{11} \text{SELFEMPL} + \beta_{12} \text{PARNAO} + \beta_{13} \text{UNEMPLOY} + \beta_{14} \text{RURAL} + \beta_{15} \text{METTILNCH} + \beta_{16} \text{LEVEL2} + \beta_{17} \text{INC74EST} + \beta_{18} \text{BLACK}.
\]

The household composition variables (NIL8, NIL14, NIL157, NIL821, M22, and F22) and the dummy variables representing education (SOMERS, COMPUS, SOMECOL, and COMPUS), occupation (SELFEMPL, PARNAO, and UNEMPLOY), urbanization (RURAL) and race (BLACK) in equation (19) are defined in Chapter 3, Table 2. The new variables introduced into this analysis are defined as follows:

- **METTILNCH**: total monthly local service charges, net of monthly extension rental fees, if any, in dollars.
- **LEVEL2**: 1 if there are two levels in the living quarters; 0 otherwise, and
- **INC74EST**: estimated 1974 total family gross income, in thousands of dollars.

Two important restrictions are inherent in the mathematical formulation of equation (19). First, the explanatory variables are entered in
Chapter 4

Although transformations of the raw data, such as logs, reciprocals, ratios, etc., could be used. In this dissertation such transformations are not considered. The second important restriction in equation (19) is that it assumes that the effect of each explanatory variable on the probability of renting an extension telephone does not depend on the value of any other independent variable (i.e., it assumes the absence of any interaction effects among the explanatory variables). The validity of this restriction will be tested in future research.

Although it is possible to use a maximum likelihood estimation procedure to obtain estimates of \( \beta \) in equation (18) assuming only that the dependent variable is an increasing or decreasing function of each explanatory variable, Warner (1962, p. 16) regards this approach as being prohibitively expensive. Computational difficulties thus require additional assumptions regarding the form of the relationship between \( \text{Pr}(\text{RET} = 1) \) and \( X_{13} \), the form for \( F \) in equation (18) must be specified. Several commonly used functional forms will now be considered.

**Linear Probability Model**

The simplest functional form for this relationship, from a computational point of view, can be obtained by assuming that \( F \) in equation (18) is the identity function. In this case equation (18) can be rewritten as

\[
\text{Pr}(\text{RET} = 1) = X_{13},
\]

and ordinary least squares regression can be used to estimate \( \beta \). Although a linear probability model will produce estimates that are unbiased and consistent (assuming the true response curve is linear) and can be obtained at low costs, there are several difficulties with this approach. The first problem is that, due to the discrete nature of the dependent variable, the error terms cannot be assumed to have equal variance. Although Zellner and Lee (1965, p. 38) suggest that this heteroscedasticity may be removed by using a generalized least squares regression procedure, Demsetz and McFadden (1975, p. 103) do not recommend this approach since it aggravates the model's sensitivity to specification error.

An additional difficulty encountered in using ordinary least squares in a dichotomous dependent variable situation is that extremely low or extremely high values of \( X_{13} \) might produce a predicted value for the dependent variable that is less than zero or greater than one. This, of course, would be inconsistent with interpreting the expected value of the dependent variable as the probability of a household renting an extension telephone. Ignoring this inconsistency and setting \( \text{Pr}(\text{RET} = 1) \) equal to one when \( X_{13} > 1 \) and \( \text{Pr}(\text{RET} = 1) \) equal to zero when \( X_{13} < 0 \) would result in predicting that an alternative will be chosen with certainty even though we may observe that it sometimes is not chosen. Restricted least squares, a procedure whereby ordinary least squares regression is used to estimate \( \beta \) subject to the constraint that \( 0 < X_{13} < 1 \), will avoid predictions outside the unit interval but, according to Nerlove and Press (1973, p. 8), when heteroscedasticity is also taken into account the programming problem becomes especially complex.

Distinct from the statistical problems inherent in the linear probability model is the theoretical difficulty that the relationship between the probability of renting an extension telephone and a linear combination of the explanatory variables may not be linear. As with many purchasing decisions, each residential subscriber can be characterized as having a threshold level of \( X_{13} \), below which an extension telephone is not rented and above which one or more are
The threshold level varies among individuals as a result of differences in personality, education, tastes, etc. Kau and Hill (1972) demonstrate that when a group of individuals is considered, the functional relationship between the probability of purchasing a good characterized by a threshold level and a linear combination of independent variables will have an ogive or sigmoid shape, such as depicted in Figure 1. With this type of response curve the effect of a change in an explanatory variable on the probability of purchasing the good is not constant; it is small for low and high values of $X'B$ and high for middle values of $X'B$. In conclusion, there are theoretical as well as statistical reasons for assuming that the true response curve is more likely to look like that in Figure 1 than a straight line.

**Binary Probit and Logit Probability Models**

All of the difficulties encountered with the linear probability model can be avoided by specifying $F$ in equation (18) to be a cumulative distribution function that is an increasing function of $X'B$, translates $X'B$ into numbers between zero and unity, and has an ogive shape. The two most commonly used cumulative distribution functions are the normal and the logistic.

When $F$ in equation (18) is taken to be the standard normal cumulative distribution function, the method is referred to as binary probit analysis. In this case equation (18) can be rewritten as

\[
\Pr(RET=1) = (2\pi)^{-1/2} \int_{-\infty}^{X'B} e^{-t^2/2} dt.
\]

Although it is possible to estimate $\beta$ in equation (21) using maximum likelihood procedures, the simultaneous equations resulting from differentiation of the natural logarithm of the likelihood function with respect to each independent variable are extremely difficult to solve.

A cumulative distribution function which is similar to the cumulative normal except at the tails and which has a simpler algebraic form is the logistic. The expression for what Domencich and McFadden (1975, p. 56) refer to as a binary logit probability model is as follows:

\[
\Pr(RET=1) = \frac{\exp X'B}{1 + \exp X'B}.
\]

This statistical model states that the probability of a household renting one or more extension telephones varies with respect to $X'B$ as a logistic cumulative distribution function having the shape depicted in Figure 1. Since the logistic and probit probability functions are essentially equivalent and since there are no a priori theoretical preferences for either one, the binary logit probability model was selected as the primary statistical technique for analyzing the demand for extension telephones because of its computational advantages.

The procedure for deriving the maximum likelihood estimators for the parameters of equation (22) will now be described. Let $X_i$, $i = 1, \ldots, n$, represent the vector of values for the independent variables for
Chapter 4

the $i^{th}$ household in the sample and assume $\Pr(RET_i = 1)$ is the probability of this household renting an extension telephone. For convenience, the observations are arranged so that $i = 1, 2, \ldots, m$ are the households renting one or more extension telephones and $i = m_1 + 1, m_1 + 2, \ldots, n$ are the ones not renting one. If each observation can be assumed to be selected independently of the others, the probability of observing the sample (i.e., the likelihood function for the sample, $L$) can then be written as

$$L = \prod_{i=1}^{m_1} \Pr(RET_i = 1) \prod_{i=m_1+1}^{n} [1 - \Pr(RET_i = 1)].$$

Substituting equation (22) into equation (23) and simplifying yields

$$L = \prod_{i=1}^{m_1} \exp[X_i\beta] \prod_{i=m_1+1}^{n} \frac{1}{1 + \exp[X_i\beta]}. \tag{24}$$

The method of maximum likelihood determines the value of $\beta$ for which the calculated probability of observing the sample is the highest. Since the natural logarithm of a function increases monotonically with the function, the values of $\beta$ that maximize the likelihood function, $L$, will also be the values of $\beta$ that maximize the natural logarithm of the likelihood function, $\text{ln}L$, which can be expressed as follows:

$$\text{ln}L = m_1 \sum_{i=1}^{m_1} X_i\beta - n \sum_{i=m_1+1}^{n} \text{ln}(1 + \exp[X_i\beta]). \tag{25}$$

The values of $\beta$ that maximize $\text{ln}L$ in equation (25) can be found by differentiating $\text{ln}L$ with respect to each of the parameters contained in the vector $\beta$, setting each derivative equal to zero, and solving the resulting system. The computer program developed by Nerlove and Press (1973) was employed to find the solution to this system of simultaneous equations, which are nonlinear in the parameters. In addition to providing maximum likelihood estimates of the parameters and asymptotic estimates of their standard errors, the Nerlove-Press program supplies the natural logarithm of the likelihood function evaluated at the maximum likelihood estimates and the predicted probability that the dependent variable is equal to one for each observation.

The maximum likelihood method described here generates estimators with desirable asymptotic properties. Under very general conditions, the maximum likelihood estimators are consistent, asymptotically unbiased, asymptotically efficient, and asymptotically normally distributed.

**Empirical Results**

Both the linear probability model and the binary logit model were estimated in the empirical analysis of the demand for residential extension telephones. Although the conclusion of the previous section is that the binary logit model is theoretically and statistically more appropriate than the linear probability model, the results of estimating both models are presented here due to interest among researchers in comparing the results obtained with these two statistical techniques in a dichotomous dependent variable situation. The results of estimating equation (20) using ordinary least squares regression are contained in Table 9, and the results of estimating equation (22) using the computer program developed by Nerlove and Press (1973) for obtaining maximum likelihood estimates are shown in Table 10, Model 1. The results for each model are presented in terms of the parameter estimates, measures of the statistical reliability of the parameter estimates and various measures of the goodness of fit. To facilitate interpretation of the estimates, the sample means and standard deviations for the explanatory variables are presented in Table 11, and the matrix of simple correla-
Before examining the estimates for each explanatory variable, the ordinary least squares regression results are given in Table 12.

Table 9. Ordinary least squares regression results for the extension telephone demand model. a

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Coefficient b</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.1722*</td>
<td>.0746</td>
</tr>
<tr>
<td>N18</td>
<td>.0237</td>
<td>.0198</td>
</tr>
<tr>
<td>N914</td>
<td>.0236</td>
<td>.0218</td>
</tr>
<tr>
<td>N1517</td>
<td>.0533</td>
<td>.0299</td>
</tr>
<tr>
<td>N1821</td>
<td>.0170</td>
<td>.0254</td>
</tr>
<tr>
<td>M22</td>
<td>.1245**</td>
<td>.0411</td>
</tr>
<tr>
<td>F22</td>
<td>.0860*</td>
<td>.0417</td>
</tr>
<tr>
<td>SOMEHS</td>
<td>.0963*</td>
<td>.0430</td>
</tr>
<tr>
<td>COMPHS</td>
<td>.1597**</td>
<td>.0418</td>
</tr>
<tr>
<td>SOMECOL</td>
<td>.3123**</td>
<td>.0533</td>
</tr>
<tr>
<td>COMPCOL</td>
<td>.3116**</td>
<td>.0548</td>
</tr>
<tr>
<td>SELFEMPL</td>
<td>-.0855</td>
<td>.0586</td>
</tr>
<tr>
<td>FARMOP</td>
<td>-.0465</td>
<td>.0546</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>.0252</td>
<td>.0998</td>
</tr>
<tr>
<td>METTISCH</td>
<td>.0297**</td>
<td>.0064</td>
</tr>
<tr>
<td>RURAL</td>
<td>-.1286**</td>
<td>.0316</td>
</tr>
<tr>
<td>LEVEL82</td>
<td>.2105**</td>
<td>.0450</td>
</tr>
<tr>
<td>INCFAEST</td>
<td>.0059**</td>
<td>.0022</td>
</tr>
<tr>
<td>BLACK</td>
<td>.1861**</td>
<td>.0372</td>
</tr>
</tbody>
</table>

R²: .1769
F: 11.9300
SEE: 204.1683

aSample size is 1018.

bIndicates significance at the 5% level and **indicates significance at the 1% level using a two-tailed t-test.

cResidual sum of squares.

Table 10. Binary logit results for the extension telephone demand models. b,c

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.2770**</td>
<td>-3.2152**</td>
<td>-2.0259**</td>
<td>-4.2820**</td>
</tr>
<tr>
<td></td>
<td>(-1.976)</td>
<td>(-1.850)</td>
<td>(-1.646)</td>
<td>(-3.147)</td>
</tr>
<tr>
<td>M18</td>
<td>-.1063</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(-.0473)</td>
<td>(-.0256)</td>
<td>(-.0227)</td>
<td>(-.0268)</td>
</tr>
<tr>
<td>N914</td>
<td>.1113</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.0927)</td>
<td>(-.0727)</td>
<td>(-.0621)</td>
<td>(-.0621)</td>
</tr>
<tr>
<td>N1517</td>
<td>.2627</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.0927)</td>
<td>(-.0727)</td>
<td>(-.0621)</td>
<td>(-.0621)</td>
</tr>
<tr>
<td>N1821</td>
<td>.0785</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.0621)</td>
<td>(-.0457)</td>
<td>(-.0457)</td>
<td>(-.0457)</td>
</tr>
<tr>
<td>M22</td>
<td>.6129**</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.1026)</td>
<td>(-.0477)</td>
<td>(-.0477)</td>
<td>(-.0477)</td>
</tr>
<tr>
<td>F22</td>
<td>.4039*</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.1024)</td>
<td>(-.0477)</td>
<td>(-.0477)</td>
<td>(-.0477)</td>
</tr>
<tr>
<td>N22</td>
<td>--</td>
<td>.5108**</td>
<td>.6846**</td>
<td>.1459</td>
</tr>
<tr>
<td></td>
<td>(-.0910)</td>
<td>(-.0925)</td>
<td>(-.1278)</td>
<td>(-.1278)</td>
</tr>
<tr>
<td>SOMEHS</td>
<td>.5037*</td>
<td>.5431*</td>
<td>.5803**</td>
<td>.7891**</td>
</tr>
<tr>
<td></td>
<td>(.1084)</td>
<td>(.1059)</td>
<td>(.1067)</td>
<td>(.1067)</td>
</tr>
<tr>
<td>COMPHS</td>
<td>.8242**</td>
<td>.8611**</td>
<td>.8599**</td>
<td>.8128**</td>
</tr>
<tr>
<td></td>
<td>(.1064)</td>
<td>(.1010)</td>
<td>(.1022)</td>
<td>(.1022)</td>
</tr>
<tr>
<td>SOMECOL</td>
<td>.1863**</td>
<td>.1863**</td>
<td>.1863**</td>
<td>.1863**</td>
</tr>
<tr>
<td></td>
<td>(.1086)</td>
<td>(.1059)</td>
<td>(.1067)</td>
<td>(.1067)</td>
</tr>
<tr>
<td>COMPCOL</td>
<td>.1955**</td>
<td>.1988**</td>
<td>.2021**</td>
<td>.2044**</td>
</tr>
<tr>
<td></td>
<td>(.1215)</td>
<td>(.1215)</td>
<td>(.1215)</td>
<td>(.1215)</td>
</tr>
<tr>
<td>SELFEMPL</td>
<td>.6314**</td>
<td>.6314**</td>
<td>.6314**</td>
<td>.6314**</td>
</tr>
<tr>
<td></td>
<td>(.1215)</td>
<td>(.1215)</td>
<td>(.1215)</td>
<td>(.1215)</td>
</tr>
<tr>
<td>FARMOP</td>
<td>.2500</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.1407)</td>
<td>(-.0655)</td>
<td>(-.0655)</td>
<td>(-.0655)</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>.1139</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>(.1023)</td>
<td>(-.0555)</td>
<td>(-.0555)</td>
<td>(-.0555)</td>
</tr>
<tr>
<td>METTISCH</td>
<td>.1288**</td>
<td>.1513**</td>
<td>--</td>
<td>.1268**</td>
</tr>
<tr>
<td></td>
<td>(.0159)</td>
<td>(.0155)</td>
<td>(-.0189)</td>
<td>(-.0189)</td>
</tr>
<tr>
<td></td>
<td>(.0134)</td>
<td>(.0134)</td>
<td>(.0134)</td>
<td>(.0134)</td>
</tr>
</tbody>
</table>
Table 10 (Continued)

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMT2</td>
<td>-.631**</td>
<td>-.600**</td>
<td>-.0652</td>
<td>-1.026**</td>
</tr>
<tr>
<td></td>
<td>(-.0771)</td>
<td>(-.0748)</td>
<td>(-.0808)</td>
<td>(-.1309)</td>
</tr>
<tr>
<td>RMT4</td>
<td>-.164**</td>
<td>-.139**</td>
<td>-.0137</td>
<td>-.2665</td>
</tr>
<tr>
<td></td>
<td>(-.1077)</td>
<td>(-.1100)</td>
<td>(-.1130)</td>
<td>(-.1498)</td>
</tr>
<tr>
<td>LEVELS2</td>
<td>1.030**</td>
<td>.997**</td>
<td>.949**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.1126)</td>
<td>(.1102)</td>
<td>(.1124)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[.2185]</td>
<td>[.2193]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUBRBDM</td>
<td>-.6703**</td>
<td>(-.0753)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-.1612)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC74EST</td>
<td>.0285**</td>
<td>.0367**</td>
<td>.0354**</td>
<td>.0280**</td>
</tr>
<tr>
<td></td>
<td>(.0056)</td>
<td>(.0047)</td>
<td>(.0047)</td>
<td>(.0063)</td>
</tr>
<tr>
<td></td>
<td>(.0071)</td>
<td>(.0074)</td>
<td>(.0085)</td>
<td>(.0067)</td>
</tr>
<tr>
<td>BLACK</td>
<td>.936**</td>
<td>1.039**</td>
<td>1.065**</td>
<td>1.373**</td>
</tr>
<tr>
<td></td>
<td>(.0941)</td>
<td>(.0889)</td>
<td>(.0886)</td>
<td>(.1254)</td>
</tr>
<tr>
<td></td>
<td>[.2264]</td>
<td>[.2510]</td>
<td>[.2563]</td>
<td>[.2303]</td>
</tr>
</tbody>
</table>

\[d\] Asymptotic standard errors are in parentheses below the coefficients.

\[b\] Significance at the 5% level and ** signifies asymptotic significance at the 1% level using a two-tailed t-test.

\[c\] Numbers in brackets below the standard errors are the effects of a unit change in each explanatory variable on Pr(RET=1), evaluated at the probability associated with the mean household of the sample.

\[d\] Sample size.

Interpreting the Empirical Results

The coefficients of the variables contained in the linear probability model in Table 9 are the ordinary least squares estimates of the effects of a unit change in each explanatory variable, ceteris paribus, on the probability of a household renting an extension telephone. The coefficient of determination ($R^2$) is provided as a measure of goodness of fit in Table 9 although it is not clear exactly what meaning it has in

---

Table 11. Sample means and standard deviations for the explanatory variables used in extension demand Model 1, Table 10, and in Table 9

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N18</td>
<td>.3969</td>
<td>.7640</td>
</tr>
<tr>
<td>N914</td>
<td>.3468</td>
<td>.7141</td>
</tr>
<tr>
<td>M1517</td>
<td>.2299</td>
<td>.5323</td>
</tr>
<tr>
<td>N1821</td>
<td>.2908</td>
<td>.6047</td>
</tr>
<tr>
<td>M22</td>
<td>.8448</td>
<td>.4973</td>
</tr>
<tr>
<td>F22</td>
<td>.2790</td>
<td>.4722</td>
</tr>
<tr>
<td>SOMHHS</td>
<td>.1660</td>
<td>.3733</td>
</tr>
<tr>
<td>COMHHS</td>
<td>.2387</td>
<td>.4265</td>
</tr>
<tr>
<td>SOMECOL</td>
<td>.1090</td>
<td>.3118</td>
</tr>
<tr>
<td>COMPCOL</td>
<td>.1189</td>
<td>.3238</td>
</tr>
<tr>
<td>SELFEMP</td>
<td>.0786</td>
<td>.2652</td>
</tr>
<tr>
<td>PARHOP</td>
<td>.0845</td>
<td>.2782</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>.2151</td>
<td>.4111</td>
</tr>
<tr>
<td>MEYTLHDG</td>
<td>8.1799</td>
<td>2.3573</td>
</tr>
<tr>
<td>RURAL</td>
<td>.6444</td>
<td>.6789</td>
</tr>
<tr>
<td>LEVELS2</td>
<td>.1228</td>
<td>.3284</td>
</tr>
<tr>
<td>INC74EST</td>
<td>11.9198</td>
<td>8.5778</td>
</tr>
<tr>
<td>BLACK</td>
<td>.2809</td>
<td>.4497</td>
</tr>
</tbody>
</table>
the present case of correlating predicted probabilities with actual binary outcomes. For a given set of values for the explanatory variables for some household, the predicted values for \( \hat{y} \) from Table 10, Model 1, can be used together with equation (22) to estimate the probability of a particular family renting an extension telephone. With the binary logit probability model, in contrast to the linear probability model, the effect of a unit change in an explanatory variable on the probability of renting an extension telephone does not depend solely upon the magnitude of the estimated \( \beta \) coefficients. With the logit model the relationship between the dependent variable and a linear combination of explanatory variables is nonlinear. Thus, the partial effects of a change in one of the explanatory variables will not be the same for all households, but will depend on each household’s original probability of renting an extension station. As can be seen from Figure 1, the effect of a change in an explanatory variable decreases as the original probability of having an extension telephone moves in either direction from .5.

An expression for the effect of a small change in the \( k \)th explanatory variable (\( \delta x_k \)) can be obtained by differentiating equation (22) with

<table>
<thead>
<tr>
<th>Variable</th>
<th>( (1) )</th>
<th>( (2) )</th>
<th>( (3) )</th>
<th>( (4) )</th>
<th>( (5) )</th>
<th>( (6) )</th>
<th>( (7) )</th>
<th>( (8) )</th>
<th>( (9) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>N18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N914</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1517</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1821</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOMPERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONMPHS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOMECOL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMPFCOL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELFCOSHFL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FARMPOP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NETFLSCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RURAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEVELS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC74EST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLACK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12 (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>( (10) )</th>
<th>( (11) )</th>
<th>( (12) )</th>
<th>( (13) )</th>
<th>( (14) )</th>
<th>( (15) )</th>
<th>( (16) )</th>
<th>( (17) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>N18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N914</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1517</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N1821</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOMPERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONMPHS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOMECOL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMPFCOL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELFCOSHFL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FARMPOP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NETFLSCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RURAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEVELS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INC74EST</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLACK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For a given set of values for the explanatory variables for some household, the predicted values for \( \hat{y} \) from Table 10, Model 1, can be used together with equation (22) to estimate the probability of a particular family renting an extension telephone. With the binary logit probability model, in contrast to the linear probability model, the effect of a unit change in an explanatory variable on the probability of renting an extension telephone does not depend solely upon the

An expression for the effect of a small change in the \( k \)th explanatory variable (\( \delta x_k \)) can be obtained by differentiating equation (22) with

Extension Stations
respect to $\chi^k$. This expression is presented in equation (26):

$$Pr(RET=1) = \delta_k \Pr(RET=1) \Pr(RET=0),$$

where $\delta_k$ is the estimated coefficient for the $k$th explanatory variable and $\Pr(RET=0) = 1 - \Pr(RET=1)$. For convenience in interpreting the magnitude of the effect of each explanatory variable in the logit probability model, as well as to aid in comparing the logit estimates with the ordinary least squares results, the effect of a unit change in each explanatory variable on the probability of renting an extension telephone, evaluated at the probability associated with the mean household in the sample, is presented in Table 10 in brackets below the standard error.

The four measures of goodness of fit supplied with the logit results in Table 10 also require explanation. The natural logarithm of the likelihood function, evaluated at its maximum, can be interpreted as a measure of the goodness of fit in the sense that it approaches zero for a perfect fit. Another way to evaluate how well a binary logit probability model approximates the observed data is to determine the percentage of correct choices predicted with a probability of 0.5 or more. A prediction is termed correct if a household actually rents an extension station and the predicted probability is greater than 0.5, or if a household does not rent an extension telephone and the predicted probability is less than 0.5.

McFadden (1976b, p. 59) suggests that a more reliable and useful measure of goodness of fit, given the maximum likelihood estimation method being used, is the likelihood ratio statistic (LRS) defined as

$$\text{LRS} = -2 \ln \left( \frac{L_2}{L_1} \right),$$

where $L_1$ is the likelihood function evaluated at the maximum likelihood estimates for the model under consideration, and $L_2$ is the likelihood function evaluated at the maximum likelihood estimates for a model obtained by setting the coefficients of all the independent variables except the constant equal to zero. In large samples, under the null hypothesis that the coefficients of all the explanatory variables are equal to zero, $-2 \ln \left( \frac{L_2}{L_1} \right)$ approximates a chi-square distribution with $(m-1)$ degrees of freedom, where $m$ is the number of parameters to be estimated, including the constant term.

With this statistic a test analogous to an F-test in a regression model can be constructed. The test rule is to reject the null hypothesis whenever $\text{LRS}$ exceeds the critical value of the chi-square statistic, calculated at the desired level of significance.

Another measure of goodness of fit for the logit model, discussed in McFadden (1973, p. 121), which is comparable to the coefficient of determination in linear regression, is the likelihood ratio index ($\rho^2$) defined as

$$\rho^2 = 1 - \frac{\ln L_2}{\ln L_1},$$

where $L_1$, $L_2$, and $L$ are as defined above. Values for the likelihood ratio index fall within the $(0,1)$ interval with higher values corresponding to closer fits. According to McFadden (1975b, p. 60), values for the likelihood ratio index are typically lower than the multiple correlation coefficients obtained in ordinary least squares regression analysis in analogous applications.

Comparing Ordinary Least Squares and Logit Results

Overall, the ordinary least squares estimates, presented in Table 9, and the binary logit estimates, presented in Table 10, Model 1, are
very similar. All the variables with coefficients significantly different from zero at the 5 percent level in Table 9 are also significant in Table 10, Model 1. The signs of the coefficients of these variables are also the same. The magnitude of the coefficients obtained using ordinary least squares regression corresponds very closely with the logit estimates of the effects of a unit change in each independent variable, evaluated for the mean household. This is not surprising since estimates produced using ordinary least squares in a dichotomous dependent variable situation are known to be reasonable when most of the observations fall on the relatively linear middle part of the sigmoid response curve. This is the case with the sample used in this analysis since 42 percent of the households have one or more extension telephones. Thus, the results, in terms of significance, magnitude, and direction of effect of the explanatory variables, are largely insensitive to the method of estimation. Only the logit estimates will be referred to in the following discussion of the results for individual variables.

Income

Looking first at the income variable (INC74EST), Table 10, Model 1, indicates that its coefficient is significantly different from zero at the 1 percent level and that a $1,000 increase in income will increase the probability of the mean household having an extension telephone by .007. The results of estimating the linear probability version of this equation using ordinary least squares regression and dummy variables for different income categories indicate that, as expected, the relationship between the probability of having an extension telephone and income is sigmoid shaped.

From equation (26) the dichotomous choice analog to the ordinary income elasticity of demand can be derived. This expression is given as follows:

$$\frac{2Pr(RET=1) - Pr(RET=0)}{INC74EST} = \beta_1 Pr(RET=0) INC74EST.$$  (29)

where $\beta_1$ represents the estimated coefficient of the income variable. From equation (29) it can be seen that the income elasticity depends on the original probability of having an extension telephone. For the probability associated with the mean household in the sample, the income elasticity of demand for extension telephones is .209. Thus, the demand for residential extension telephones, as was found to be the case with toll calls, is income inelastic. This estimate is slightly larger than the income elasticity of .107 calculated by Irish (1974, p. 125, see footnote 12) using pooled cross-section and time-series data for individual exchanges in North Carolina.

Traditional Noneconomic Variables

None of the dummy variables representing occupation of the head of the family in Model 1, Table 10, has a coefficient that is significantly different from zero at the 10 percent level. In contrast to the demand for long-distance calls, there are no strong a priori theoretical reasons for expecting occupation to affect the demand for extension telephones, other than the effect it might exert on tastes and preferences.

Although all the variables representing family size and composition in Table 10, Model 1, have the correct sign, only the coefficients for males and females 22 years of age or older (M22 and F22) are significantly different from zero at the 5 percent level. This is not surprising since younger family members are expected to have a smaller
specific effect (i.e., need) for extension telephones than adult members of the household.

In order to test the null hypothesis that the coefficients of $M2$ and $P2$ are equal, and, at the same time, test the hypothesis that all the occupation dummy variables (SELFEMPL, FARMOP, and UNEMPLOY) and all the other household composition variables ($N18$, $N914$, $N1517$, and $NI821$) have coefficients equal to zero, a likelihood ratio test was performed. To perform this test binary logit analysis was used to re-estimate Model 1, Table 10, after deleting all the occupation dummy variables and household composition variables and adding $N22$, the number of household members 22 years of age or older ($N22 = F22 + M22$). These results are presented in Table 10, Model 2. The test is based on the fact that, under the null hypothesis, $-2$ times the natural logarithm of the ratio of the likelihood functions of the two models, which is equal to 12.893, is approximated by a chi-square distribution with 8 degrees of freedom. This test shows that the null hypothesis cannot be rejected at the 95 percent confidence level. Model 2, Table 10, thus contains the preferred specification for household composition and reveals that an additional family member 22 years of age or older increases the mean family's probability of renting an extension telephone by 0.12.

Turning now to the remaining traditional taste and preference variables, Models 1 and 2 of Table 10 indicate that race, education, and urbanization exert a highly significant effect on the probability of renting an extension telephone. As hypothesized in the theoretical section, blacks are indeed more likely to have an extension telephone than whites. As with the demand for toll calls, the magnitude of this effect is surprisingly large. Model 2 indicates that, at the probability associated with the mean household, a black family has a 0.25 higher probability of renting an extension telephone than a similarly situated white family.

The estimated coefficients for the education dummy variables reveal that the probability of renting an extension telephone increases with the educational level attained by the head of the household. This direct relationship may be due to education serving as a proxy for permanent income or it may simply reflect a taste difference. Due to the high costs of performing logit analysis, a likelihood ratio test of the joint hypothesis that the coefficients of $SOMERS$ and $COMPS$ are equal and that the coefficients of $SOMECOL$ and $COMPCOL$ are equal was not conducted. This hypothesis could not, however, be rejected at the 5 percent level of significance using an $F$-test and ordinary least squares estimates of the linear probability version of the model.

An $F$-test and ordinary least squares regression did reveal a significant (at the one percent level) difference in the probability of renting an extension telephone between families whose head had attended or completed high school and households whose head had attended or completed college. It is expected that similar conclusions would be reached using the likelihood ratio test and logit analysis.

The logit results presented in Table 10 suggest that households located in rural areas have a significantly smaller probability (0.15 for the mean household) of renting an extension telephone than urban households. As in the long-distance demand analysis, dummy variables for the two rural subdivisions, rural farm (RF) and rural nonfarm (RNF), could not be included in the same equation containing $FARMOP$. However, when ordinary least squares was used to re-estimate Model 2, which does not contain $FARMOP$, using RF and RN dummy variables instead of RURAL, RF had a larger, in absolute value, negative coefficient than RN, although an $F$-test revealed that the difference was not significant at
the 5 percent level. Thus, Model 2 incorporates the appropriate specification for the urbanization variable.

Special Noneconomic Variables

The two noneconomic variables specific to the service under analysis remain to be discussed. Table 10 shows that NETTLSCH, a proxy variable for the amount that the family uses the telephone for local service, and LEVELS2, a proxy variable for housing unit size, exert the theoretically anticipated direction of effect on the probability of renting an extension station and have coefficients that are significantly different from zero at the 1 percent level. Having two levels in a housing unit rather than one, ceteris paribus, increases the probability of renting an extension by .24 for the mean household. An additional dollar spent on local services by the mean household increases the probability of renting an extension telephone by .04.

An alternative measure of local service usage, which was employed in the toll demand analysis, is grade of main station service. The hypothesis is that those subscribers renting a higher quality grade of service are more likely to be high local service users and thus have a greater probability of renting an extension telephone. The binary logit results of re-estimating Model 2 in Table 10 after deleting NETTLSCH and adding dummy variables representing two-party service (RMT2) and four-party service (RHT4) are presented in Table 10, Model 3. The coefficient of each of these variables has the theoretically expected sign and is significantly different from zero at the 1 percent level. Although these estimates indicate that a household receiving four-party telephone service is less likely to have an extension telephone than a similar family with a two-party line, an F-test performed using ordinary least squares regression could not reject (at the 5 percent level) the null hypothesis that the coefficients of these two dummy variables are equal.

An alternative specification was also employed for size of housing unit. For the subset of households that were also interviewed in 1973 and that owned or were buying their home at that time and did not move between the 1973 and 1975 interviews, Model 2, Table 10, was re-estimated using the number of bedrooms in the housing unit (NOBEDRM) instead of LEVELS2.92 These binary logit estimates, presented in Table 10, Model 4, suggest that an additional bedroom, ceteris paribus, increases the probability that the mean household will rent an extension telephone by .16. Due to the high degree of correlation between LEVELS2 and NOBEDRM, both of these variables are not significantly different from zero when included in the same equation. The goodness-of-fit measures at the bottom of Table 10 suggest that number of bedrooms may be a better measure of housing unit size than number of levels of living quarters.

Goodness of Fit

In spite of the many significant variables in Table 10, all of the goodness-of-fit measures provided reveal that the theoretical model of the demand for residential extension stations is not fitting the observed data very well. The natural logarithms of the likelihood functions are large in absolute value, and the likelihood ratio indices are between .13 and .18. Model 2 is capable of predicting the correct choice for only 67.68 percent of the households included in the sample. However, the likelihood ratio statistic for this model indicates that the null hypothesis that the coefficients of all the explanatory variables are equal to zero can be rejected at the 99 percent confidence level.

The poor fit achieved with the residential extension demand model...
results both from dealing with data on a large number of individual households, rather than with aggregate data, and from studying a problem that by nature is subject to many unsystematic influences. However, as long as the variables not included in the model are not correlated with those that are, unbiased estimates can still be obtained.

5

Demand for Residential Main Stations

In this chapter the factors affecting a family's decision to rent a residential main station and its choice among the available grades of service (i.e., one-party, two-party, or four-party) are analyzed theoretically and empirically. For those households with only one grade of service available, a binary logit model is developed to explain their choice between renting a one-party main station and not subscribing to telephone service. A series of binary logit probability models are constructed to simulate the decision-making process by households with more than one grade of service available. Each theoretical model contains economic and noneconomic variables and telephone system characteristics hypothesized to affect the binary choice. Since the dependent variable in each of these models is qualitative, binary logit
analysis is the primary statistical technique employed in analyzing data on subsets of households interviewed in the Second Survey of Labor Supply in Eastern North Carolina.

Theoretical Models

The grades of residential main station telephone service available to a household located within the Carolina Telephone and Telegraph Company service area depend on the section of the exchange in which it is located. The only grade of service available to families living within the base rate area of an exchange is one-party or individual line service. Households residing in mileage zones 1A and 1B may choose between one-party and two-party main station service. Households located in mileage zones 2A through 11 may select either one-party, two-party, or four-party exchange service. Of course, all households have the option of not receiving any telephone service.

Since households located within a base rate area have the choice between not subscribing to telephone service and renting a one-party main station, the binary logit probability model, described in detail in Chapter 4, is the appropriate statistical model for analyzing this dichotomous choice. The choice faced by households living in mileage zones 1A and 1B among not receiving telephone service, renting a one-party line, or renting a two-party main station is not analyzed in this dissertation. Only 120 households in the sample were located in this section of an exchange. This is considered to be an inadequate number of observations for obtaining reliable parameter estimates for the large number of variables that would be used to model this choice.

Although households living outside mileage zone 1B must choose among four alternatives (i.e., no telephone, one-party, two-party, or four-party main station), this problem can still be formulated as a dichotomous decision by assuming that the household follows a tree decision structure where at each decision point the household is faced with only a choice between two alternatives. Households located in this section of an exchange are hypothesized to follow a three-stage decision sequence. In the first stage, the household considers all the grades of service available and decides whether or not to rent a main station. In the second stage, those households that have elected to receive telephone service choose between one-party service and multi-party service. Households selecting multi-party service must then, in stage three, choose either two-party or four-party service. This sequential decision-making process is illustrated with the use of a diagram in Figure 2. This behavioral model not only simulates the household's real choice procedure but also simplifies the empirical analysis since binary logit analysis can be used to analyze each of the three dichotomous choices.

![Figure 2. Sequential decision-making process for households located outside mileage zone 1B.](image-url)
An alternative method of analyzing choices among more than two alternatives is to assume that the household considers all options open to it simultaneously. McFadden (1973) has developed a generalized form of logit analysis, variously referred to as conditional, multinomial, multiple choice, or polychotomous logit analysis, where the probability that the \( k \)th household will select the \( j \)th alternative from the set of \( J_k \) alternatives open to it is assumed to be determined by the equation

\[
P_{kj} = \frac{e^{\sum_i x_{ki} b_i}}{e^{\sum_i x_{ki} b_i} + \sum_{j'} e^{\sum_i x_{k j'} b_i}},
\]

where \( x_{ki} \) is a column vector of explanatory variables for the \( k \)th household for the \( i \)th alternative and \( b_i \) is a column vector of parameters for the \( i \)th alternative to be estimated. With data for a sample of households on the actual choices made and values for the explanatory variables for each available alternative, maximum likelihood estimates of the parameters in equation (30) can be obtained using the XLOGIT program of Willis.

Polychotomous logit analysis has been used successfully in other economic applications and is the ideal statistical technique to use in analyzing a household's decision to subscribe to telephone service and choice among available grades of service. With conditional logit analysis the number of alternatives facing each observation need not be equal so that a separate analysis for households living inside a base rate area and for households located outside mileage zone 1B would not be necessary. Also, with this statistical model the explanatory variables assumed to influence the choice of grade of service need not be the same factors assumed to affect the decision to rent a main station.

Although the XLOGIT program was available for use, multiple choice logit analysis was not employed in the main station demand analysis for the following three reasons. First, there is a limit on the number of independent variables the XLOGIT program can handle, which is exceeded by the original formulation of the model. Secondly, the XLOGIT program is very expensive to use. Thirdly, Johnson (1977, see footnote 102), in a study of college-going decisions, concluded informally that a series of dichotomous choice models, estimated using binary logit analysis, provided a superior fit to the data than the polychotomous decision model, estimated with the XLOGIT program. Similar results are anticipated with the problem under analysis since the sequential decision-making process postulated here seems to be a good simulation of the real choice procedure.

A probability choice model, containing economic variables, taste and preference variables, and telephone system attributes, is developed in the following pages to explain each of the above-discussed dichotomous choices. The first two models contain those factors hypothesized to influence a household's decision to rent a main telephone rather than going without telephone service. This choice is modeled separately for households located within a base rate area, who, for convenience, are referred to as urban households, and for families living outside mileage zone 1B, who are referred to as rural households. The third model formulated explains the determinants affecting a rural telephone subscriber's choice between one-party and multi-party service. For those households selecting multi-party service, the fourth model explains their choice between two-party and four-party service.

Phone Versus No Phone

In the hypothesized sequential decision-making process, the first decision that each household must make is whether or not it is going
to subscribe to telephone service. A telephone rented from the telephone company for a monthly fee provides the renter with access to both the local and long-distance telephone networks. The monthly flat rate entitles a subscriber to an unlimited number of local calls. There are no usage sensitive pricing options available. Demand theory and previous main station demand studies, particularly that of Perl (1975, see footnote 14), suggest that family income, monthly rental rate for telephone service, noneconomic variables such as household size and composition, race and social class, and a telephone system attribute, calling scope, may be important factors affecting the probability of a household renting a main telephone, regardless of the section of the exchange in which the household is located. The measurement of each of these explanatory variables and their expected direction of effect on the probability of a household receiving telephone service will now be considered.

Income. Family income, as measured by total family 1974 gross income, is expected to exert a positive effect on the probability of a household receiving telephone service. All of the difficulties encountered in using current measured income in the toll demand analysis in Chapter 3 are applicable to this problem.

Prices. The effects of price on the demand for toll calls and extension telephones could not be analyzed empirically due to a lack of cross-sectional variation in the prices of these two services. This, fortunately, is not the case with the price of main stations. The price of a particular grade of service is not the same for all households in the sample. Telephone rates depend on the size of the calling scope of the exchange in which the household is located as well as upon the mileage zone in which the family resides. Due to the "value of service" pricing policy adhered to by the

<table>
<thead>
<tr>
<th>Calling scope group</th>
<th>Residential main station monthly base rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One-party line</td>
</tr>
<tr>
<td>0-1,000</td>
<td>$5.25</td>
</tr>
<tr>
<td>1,001-2,000</td>
<td>5.50</td>
</tr>
<tr>
<td>2,001-4,000</td>
<td>6.00</td>
</tr>
<tr>
<td>4,001-7,500</td>
<td>6.50</td>
</tr>
<tr>
<td>7,501-12,500</td>
<td>7.00</td>
</tr>
<tr>
<td>12,501-20,000</td>
<td>7.25</td>
</tr>
<tr>
<td>20,001-32,000</td>
<td>7.25</td>
</tr>
<tr>
<td>32,001-50,000</td>
<td>7.50</td>
</tr>
</tbody>
</table>

Table 14. Carolina Telephone and Telegraph Company mileage zone charges

<table>
<thead>
<tr>
<th>Mileage zone</th>
<th>Residential zone charges, per month</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One-party line</td>
</tr>
<tr>
<td>1A</td>
<td>$ .60</td>
</tr>
<tr>
<td>1B</td>
<td>1.50</td>
</tr>
<tr>
<td>2A</td>
<td>3.00</td>
</tr>
<tr>
<td>2B</td>
<td>4.50</td>
</tr>
<tr>
<td>3</td>
<td>6.25</td>
</tr>
<tr>
<td>4</td>
<td>8.00</td>
</tr>
<tr>
<td>5</td>
<td>9.75</td>
</tr>
<tr>
<td>6</td>
<td>11.50</td>
</tr>
<tr>
<td>7</td>
<td>13.25</td>
</tr>
<tr>
<td>8</td>
<td>15.00</td>
</tr>
<tr>
<td>9</td>
<td>16.75</td>
</tr>
<tr>
<td>10</td>
<td>18.50</td>
</tr>
<tr>
<td>11</td>
<td>20.25</td>
</tr>
</tbody>
</table>


Mileage zone charges apply in connection with main stations which are located outside the base rate area, but within the exchange area, and are in addition to the base rate for the class of service furnished.

Renting a main telephone are expected to be inversely related. Households located outside mileage zone 1B may select from among three grades of service, whereas only one-party service is available to families living within a base rate area. Since the price of telephone service is not the same for the two groups, a separate estimating equation is developed for each of them.

The appropriate price facing urban households is the monthly base rate for a one-party residential main station, which is the only grade available. The monthly rental rates for all three grades of service were included in the preliminary estimating equations for rural households. However, attempts at estimating linear probability models using ordinary least squares regression indicated that inclusion of all three prices produces almost complete multicollinearity. Therefore, only the monthly rental rate for two-party residential main station telephone service is included in the model to represent the price of telephone service for households located outside mileage zone 1B.

This single price variable is assumed to capture the whole structure of rates facing the rural consumer.

A pay phone may be considered a substitute for renting a main station from the telephone company. Although, as Kotowit and Waverman point out, it may be a good substitute for making outgoing long-distance calls, it is not likely to be a good substitute for receiving incoming calls or making outgoing local calls. In any event, the price of a pay phone call, 10 cents in 1975, is the same for all households in the sample. Thus, its effect on the demand for main stations cannot be investigated.

If long-distance calling can be regarded as a complement to a residential main station, then the price structure for toll calls may be inversely related to the probability of subscribing to telephone service. The effect of toll rates on the decision to rent a main telephone is not expected to be important since a pay phone is a good substitute for making outgoing toll calls, and it seems reasonable to assume that most consumers rent a telephone primarily to make and receive local calls. Since the long-distance rate structure is the same for all households in the sample, the effects of toll rates on the decision to receive telephone service cannot be empirically examined.

Non-economic Variables. Additional explanatory variables representing household size and composition, race, and education and occupation of

Main Stations

123
the head of the household are introduced into the model to take account
of differences in tastes and preferences among the households sampled.\textsuperscript{107}

The household composition categories used in modeling the demand for main
stations are slightly different from those used in the toll and
extension demand analyses. These changes were made in order to test
the hypothesis put forth by Perl (1975, p. 6, see footnote 14) that
age reduces motility and increases reliance on the telephone as a means
of communication. The following age groups are used to represent house­
hold composition: children 8 years of age or younger, children between
the ages of 9 and 16, adolescents between the ages of 15 and 17, adults
between the ages of 18 and 21, adults between the ages of 22 and 39, and
adults 60 years of age or older. Assuming the income effect does not
outweigh the specific effect, an additional family member in each of
these age categories, \textit{ceteris paribus}, is expected to increase the
probability of the household renting a telephone. The magnitude of
this effect is expected to be smaller for the younger age groups since
their need for communication services is probably not as great as that
of adults.

The expected direction of effect of the other taste variables (e.g.,
race, education, and occupation) cannot be predicted from utility theory
alone. In an ordinary least squares multiple regression analysis of
Census data for individual households, Perl (1975, see footnote 14) found
blacks less likely to have a telephone available than whites and dis­
covered that education of the head is positively related to the proba­
bility of having a telephone available. Similar results are anticipated
in this study. Families headed by farm operators or self-employed
individuals may, \textit{a priori}, be expected to have a greater need for
telephone service, and hence a greater probability of renting a main
telephone, than families headed by a wage-salary worker.\textsuperscript{108}

Telephone company officials report observing many farm workers having telephones
installed in the spring, when they are working, and disconnected in the
fall, when they are not. Thus, families with an unemployed head may have
a lower probability of renting a main telephone than other households.

An attribute of the telephone system, calling scope, is also
included in the theoretical model to take account of differences in the
number of telephones a household may call without a toll charge.\textsuperscript{109} The
larger the calling scope of the exchange in which a household is
located, the greater is the utility derived from telephone service, and,
\textit{ceteris paribus}, the greater is the probability of renting a tele­
phone.\textsuperscript{110}

\textbf{Omitted Variables.} A variable not included in the theoretical
model developed above, but one which telephone company officials
believe may be an important factor affecting whether or not a household
is renting a main telephone, is credit rating. Under North Carolina
Utilities Commission Rule R12-1, an applicant for telephone service may
be required to pay an advance payment and a deposit before telephone
service is established. The amount required depends on the credit risk
of the individual.\textsuperscript{111}

The large amount of advance payments and deposits requested from
individuals with poor or unknown credit ratings may be an important
factor prohibiting these people from receiving telephone service.
Since these people have a poor credit rating, they probably have few
assets and may be unable to borrow the amount requested. Credit informa­
tion is unavailable for all the households surveyed making it impossible
to test this hypothesis empirically.

The effects of the service connection charge that individuals with
good credit ratings must pay with their first month’s bill also cannot
be investigated since it is the same amount, $15, for all households
Chapter 5

Interviewed. The service connection charge is not considered to be an important factor prohibiting households with good credit ratings from receiving telephone service. Individuals with good credit ratings presumably would not need to borrow money to have a telephone installed, and, if they did, they would be capable of borrowing this amount and repaying it over the time period that the telephone is rented.

Selection of the Statistical Model. With the available disaggregated data, observations on residential main telephones are (0,1) data rather than quantities of residential main stations. The natural method of analysis is thus a probabilistic model of choice behavior. Due to the statistical problems inherent in a linear probability model, discussed in the previous chapter, and the expectation that the relationship between a linear combination of independent variables and the probability of renting a telephone is a sigmoid-shaped curve, the binary logit probability model is the appropriate statistical model. Letting PHONE-1 if a household rents a main telephone, and letting PHONE-0 if it does not, the above-developed theory can be expressed in the form of a binary logit probability model as follows:

\[
\Pr (\text{PHONE}=1) = \frac{e^{Y'\beta}}{1 + e^{Y'\beta}},
\]

where \( Y \) is a column vector of values for the explanatory variables for the household and \( \beta \) is a column vector of parameters to be estimated such that

\[
Y'\beta = \beta_0 + \beta_1 \text{INC74EST} + \beta_2 \text{MIS} + \beta_3 \text{NSIV1} + \beta_4 \text{NSIV2} + \beta_5 \text{PRICE} + \beta_6 \text{SOCHEX} + \beta_7 \text{SHEC} + \beta_8 \text{SELPEHPL} + \beta_9 \text{FARMOP} + \beta_{10} \text{UNEMPLOY} + \beta_{11} \text{BLACK} + \beta_{12} \text{CSAVG}.
\]

In equation (32), \( \text{PRICE} = \text{PRHT1} \) for households living within a base rate area and \( \text{PRICE} = \text{PRHT2} \) for households located outside mileage zone 1B. All of the variables contained in equation (32) are defined in Table 15.

There are two restrictions inherent in the above developed statistical model for the phone versus no phone choice. The first restriction is that the model assumes that all independent variables enter equation (32) in a linear form. The variables in this function may be transformations of the original data, such as logs, reciprocals, ratios, etc. Such transformations were not considered in this research. A more important restriction in this model is the assumption that the effects of the explanatory variables on the decision to rent a telephone are independent of each other (i.e., that there are no interactions among the independent variables). This may not be a valid assumption. The effect of price of telephone service on the probability of renting a main station may depend on a household's income. One would expect the decrease in the probability of renting a telephone due to an increase in price to be greater for low income families than for high income people. The results of a preliminary test for such an interaction are discussed in the empirical section of this chapter.

One-Party versus Multi-Party Service

Households located in mileage zones 2A through 11 that have elected to receive telephone service (i.e., that have reached decision point 2 in Figure 2) hypothetically now choose between an individual line and a multi-party line (i.e., a two- or four-party line). This choice is assumed to be dependent upon family income, price of each alternative, tastes and preferences (i.e., household size and composition, race and occupation and education of the head of the household),
and two telephone system attributes, the average number of hookups on
two-party and on four-party lines for the exchange in which the

Table 15. Definition of the explanatory variables used in the main
station demand model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>INC74EST</td>
<td>Estimated total household 1974 gross income, in thousands of dollars.</td>
</tr>
<tr>
<td>PRMT2, PRMT2, PRMT4</td>
<td>Monthly rental rate (including mileage zone charges) for 1-, 2-, and 4-party residential main station telephone service, respectively.</td>
</tr>
<tr>
<td>PRMT2AV</td>
<td>Arithmetic average of PRMT2 and PRMT4.</td>
</tr>
<tr>
<td>N8, N9, N10, N11, N12, N13, N14</td>
<td>Number of household members 8 years of age or younger, between the ages of 9 and 14, between the ages of 15 and 17, between the ages of 18 and 21, and between the ages of 22 and 59, respectively.</td>
</tr>
<tr>
<td>N22</td>
<td>Number of household members 22 years of age or older.</td>
</tr>
<tr>
<td>N60</td>
<td>Number of household members 60 years of age or older.</td>
</tr>
<tr>
<td>SCHOOLS</td>
<td>1 if family head completed more than 8 but less than 12 years of school; 0 otherwise.</td>
</tr>
<tr>
<td>COMPFS</td>
<td>1 if highest grade in school completed by head was the twelfth; 0 otherwise.</td>
</tr>
<tr>
<td>SODESCHOOL</td>
<td>1 if head completed grade more than 12 but less than 16 years of education; 0 otherwise.</td>
</tr>
<tr>
<td>COMPCOL</td>
<td>1 if head completed 16 or more years of schooling; 0 otherwise.</td>
</tr>
<tr>
<td>COL</td>
<td>1 if head completed more than 12 years of schooling; 0 otherwise.</td>
</tr>
<tr>
<td>SELFEMPL</td>
<td>1 if head is self-employed; 0 otherwise.</td>
</tr>
<tr>
<td>FARMOP</td>
<td>1 if head is a farm operator; 0 otherwise.</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>1 if head did not earn any income during 1974; 0 otherwise.</td>
</tr>
<tr>
<td>BLACK</td>
<td>1 if family is black; 0 if white.</td>
</tr>
<tr>
<td>CSAVG</td>
<td>Average of February, March, April, and May 1975 calling scope for the exchange in which the household is located, in thousands.</td>
</tr>
<tr>
<td>AVGMT2, AVGMT4</td>
<td>Average number of hookups on two-party and four-party lines, respectively, on May 31, 1975, for the exchange in which the household is located.</td>
</tr>
</tbody>
</table>

Prices for all three grades of service cannot be included in an estimating equation without introducing severe multicollinearity. Therefore, the arithmetic average of two-party and four-party service was selected to represent the price of multi-party service. This variable is introduced into the model along with the monthly rental rate for a one-party main station. Demand theory predicts that an increase in the price of multi-party service, \textit{ceteris paribus}, should increase a household's probability of selecting individual line service, whereas an increase in the price of one-party service, \textit{ceteris paribus}, should decrease this probability.

An increase in family income or in the size of the household should, \textit{ceteris paribus}, increase the probability of selecting a one-party main station. Fewer age groups are used in this model to specify household composition than were used in the phone versus no phone model due to a desire to keep the number of independent variables as small as possible because of the smaller sample size. Also, the age composition of the household is not believed to be an important factor affecting this choice. The signs on the dummy variables representing race, education and occupation of the family head cannot be predicted \textit{a priori}.

The average number of hookups on two-party and on four-party lines for the exchange in which the subscriber is located are estimates of the total number of subscribers a residential telephone customer would expect to find on his party line if he selected two- or four-party service, respectively. It is hypothesized that the more hookups a subscriber expects there to be on two- and four-party lines, everything else being the same, the more likely he will be to select a private
line. For example, if a family thinks they could rent a two-party main station with no one else on their line, they would be more likely to do this than to pay more for a one-party main station and receive almost the same quality of service.

Letting $RMT1 = 1$ if a telephone subscriber located outside mileage zone 1B rents a one-party main station and letting $RMT1 = 0$ otherwise, the conditional probability that a household electing to receive telephone service will select one-party rather than multi-party service can be expressed in the form of a binary logit probability model as follows:

$$
Pr(RMT1 = 1) = \frac{\exp Z'y}{1 + \exp Z'y},
$$

where $Z$ is a column vector of values for the explanatory variables for the household and $y$ is a column vector of parameters to be estimated such that:

$$Z'y = \gamma_0 + \gamma_1 \text{INC74EST} + \gamma_2 \text{N18} + \gamma_3 \text{N19} + \gamma_4 \text{N117} + \gamma_5 \text{N121} + \gamma_6 \text{N22} + \gamma_7 \text{PRMT2} + \gamma_8 \text{SOMERS} + \gamma_9 \text{COMPS} + \gamma_{10} \text{COL} + \gamma_{11} \text{SELFEMPL} + \gamma_{12} \text{FARMOP} + \gamma_{13} \text{BLACK} + \gamma_{14} \text{AVGHKHT2} + \gamma_{15} \text{AVGHKHT4}.
$$

The variables contained in equation (34) are defined in Table 15. All the explanatory variables are entered linearly in this equation and the absence of interactions among the explanatory variables is again assumed.

**Two-Party versus Four-Party Service**

Those families living outside mileage zone 1B that elect to receive multi-party service instead of renting a one-party main station (i.e., those households that have reached decision point 4 in Figure 2) must next select either a two-party or a four-party line. Demand theory suggests that this decision will be influenced by household income, household size and composition, price of each alternative, race, education and occupation of the head of the household, and the average number of hookups on four-party lines for the exchange in which the household is located. The anticipated direction of effect of each of these variables on the probability of renting a two-party main station will now be considered.

The household's estimated total 1974 gross income, household size, monthly rental rate for a four-party main station, and the average number of hookups on four-party lines, as defined previously, are expected to have a positive effect on the probability of renting a two-party line. The price of two-party service is expected to exert a negative influence on this probability. The effects of the dummy variables representing tastes (i.e., occupation, education, and race) cannot be predicted a priori.

Letting $RMT2 = 1$ if a multi-party telephone subscriber rents a two-party main station and letting $RMT2 = 0$ otherwise, the conditional probability that a household choosing multi-party service selects two-party rather than four-party service can be expressed in the form of a binary logit probability model as follows:

$$
Pr(RMT2 = 1) = \frac{\exp A'y}{1 + \exp A'y},
$$

where $A$ is a column vector of values for the independent variables for the household and $y$ is a column vector of parameters to be estimated such that:

$$A'y = \delta_0 + \delta_1 \text{INC74EST} + \delta_2 \text{N18} + \delta_3 \text{N22} + \delta_4 \text{N117} + \delta_5 \text{N121} + \delta_6 \text{N22} + \delta_7 \text{PRMT2} + \delta_8 \text{PRMT4} + \delta_9 \text{SOMERS} + \delta_{10} \text{COMPS} + \delta_{11} \text{COL} + \delta_{12} \text{SELFEMPL} + \delta_{13} \text{FARMOP} + \delta_{14} \text{UNEMPLOY} + \delta_{15} \text{BLACK} + \delta_{16} \text{AVGHKHT4}.
$$
The variables included in equation (36) are defined in Table 15. As in the other models, all variables are entered linearly in equation (36), and it is assumed that there are no interactions among the explanatory variables.

Data

Data on subsets of households interviewed in the Second Survey of Labor Supply in Eastern North Carolina were used to estimate each of the four binary logit probability models developed in the preceding sections of this chapter. The observations used in each estimation will now be described.

Data on households located within the base rate area of a CT&T exchange and within the area served by the Tri-County Telephone Membership Corporation, which offers only one-party main station service and does not assess mileage charges, were used to estimate the parameters of the phone versus no phone model for urban households (i.e., equation (11) with PRICE = PRMT1). The dependent variable in this equation was set equal to unity for those households renting a main station on the day of the interview and was set equal to zero for those that were not. Since two-party residential main station telephone service has been an obsolete service offering in CT&T base rate areas since March 5, 1973, some of the households reporting a telephone installed on the day of the interview might actually have been renting a two-party line. The actual grade of service is known only for those households giving written permission for CT&T to release their telephone records. Of those households the local service records indicated that 18 were renting two-party lines. In order to maintain consistency with the price variable used in this model, the monthly rental rate for a one-party main station, these households were deleted from the data set.

Observations were also deleted if these records indicated that the telephone was a business main station, that someone else was paying for the telephone service, or that a member of the household was a telephone company employee and thus entitled to a concession. All of the urban households not signing the permit slip, but reporting a telephone installed in their housing unit, were assumed to be renting a one-party residential main station. Additional observations were deleted if data were missing for any of the explanatory variables included in the model. These deletions left 924 households, of which 80.92 percent reported a telephone installed in their home on the day interviewed.

The binary logit probability model for the phone versus no phone choice for rural households (i.e., equation (31) with PRICE = PRMT1) was estimated using data on households located in CT&T mileage zones 1A through 11. The dependent variable in this equation was set equal to unity for families renting a main telephone on the day of the interview and was set equal to zero for those that were not. Observations on permit-signing households were deleted if telephone company records indicated that the telephone installed in the housing unit was classified as a business phone, that someone else was paying the telephone bill, or that a member of the family was working for the telephone company. Additional observations were deleted if data on any of the independent variables were missing. Of the 863 households remaining, 78.91 percent reported having a telephone installed in their housing unit at the time of the interview.

The binary logit probability model for the multi-party versus one-party grade of service choice (i.e., equation (33)) was estimated using data on CT&T residential subscribers living outside mileage zone 1B who granted permission for information on their grade of service to be
released by the telephone company. For those rural telephone subscribers renting a one-party main station on June 20, 1975, the day the grade of service information was obtained from CT&T, the dependent variable in this equation was set equal to unity; otherwise it was set equal to zero. Of the 464 households included in this sample, 31.90 percent were receiving individual line service at the time the grade of service information was obtained on June 20, 1975.

The binary logit probability model for the two-party versus four-party grade of service choice (i.e., equation (35)) was estimated using data on permit-signing, multi-party residential subscribers located outside mileage zone 18. The dependent variable in this equation was set equal to unity for multi-party residential subscribers who were renting a two-party main station on June 20, 1975; otherwise it was set equal to zero. Of the 316 households included in this sample, 28.80 percent were renting a two-party line on the day when the grade of service information was obtained.

Since the aim of this analysis of the demand for main stations is to estimate the parameters of each of the four models discussed above, and not to prepare estimates of a variable's mean or proportion for some aggregate of households, the randomness or representativeness of each of the samples is not considered vital. As mentioned previously, even if a sample is not random, parameter estimates will not be biased as long as there is no interaction between the explanatory variables and variables related to the probability of being included in the sample. When logit analysis is the statistical technique employed, a random sample is useful for forecasting purposes, as will be demonstrated in the following chapter. To determine if interactions exist and to assess the appropriateness of using the sample for forecasting, the randomness of the sample used in the main station demand analysis was evaluated by comparing information that was available from the survey with information obtained from independent sources.

The percentage of telephone subscribers in the sample renting 1-party, 2-party, and 4-party main stations is within a 95 percent confidence interval of the corresponding percentage of residential telephone subscribers renting each of these grades of service, as calculated from the May 31, 1975, CT&T station development reports for the exchanges located within the eight-county survey area. Thus, it can be assumed that the sample is random with respect to choice among grades of service.

A comparison of the percentage of households with a telephone installed, calculated for the sample used in this analysis, with the percentage of occupied housing units within the eight-county survey area reporting a telephone available in the 1970 Census of Housing conducted by the U. S. Bureau of the Census (1972a) reveals that 13 percent more households had a telephone available at the time of the interview than at the time of the Census. Although there are differences in definitions and sampling techniques, the major cause of this discrepancy is believed to be the timing of the surveys. The percentage of households with a telephone available has been increasing steadily over time, and one would expect it to be higher in 1975 than 1970. Also, the Census percentage of households with a telephone available is consistently lower, county by county, than the corresponding county percentages calculated using the survey data. Thus, this evidence cannot be interpreted as meaning that the sample is not random with respect to telephone availability.

Empirical Results

Both ordinary least squares regression and binary logit analysis were used in the empirical analysis of the demand for telephone service.
and choice of grade of service. The results obtained using both of these methods are presented in order to assess the effects of choice of statistical technique on the parameter estimates obtained in a dichotomous dependent variable situation. The results obtained using ordinary least squares regression to estimate the linear probability version of each of the four models developed above are presented in Table 16. The results of estimating each of the binary logit probability models using the computer program developed by Nerlove and Press (1973) for

Table 16. Ordinary least squares regression results for the main station demand models.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Base rate area</th>
<th>Mileage zone greater than LB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phone vs no phone</td>
<td>Phone vs no phone</td>
</tr>
<tr>
<td>Intercept</td>
<td>.5523*** (.1547)</td>
<td>.6312*** (.1025)</td>
</tr>
<tr>
<td>INC74EST</td>
<td>.0058*** (.0021)</td>
<td>.0013*** (.0020)</td>
</tr>
<tr>
<td>PMT1</td>
<td>.0183 (.0212)</td>
<td>--</td>
</tr>
<tr>
<td>PMT2</td>
<td>--</td>
<td>-.0114 (.0092)</td>
</tr>
<tr>
<td>PMT24AV</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>PMT74</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>N18</td>
<td>-.0432*** (.0170)</td>
<td>-.0464*** (.0258)</td>
</tr>
<tr>
<td>N914</td>
<td>-.0067 (.0178)</td>
<td>.0200 (.0168)</td>
</tr>
<tr>
<td>N1537</td>
<td>.0243 (.0266)</td>
<td>.0334 (.0260)</td>
</tr>
<tr>
<td>N1521</td>
<td>-.0265 (.0224)</td>
<td>.0463*** (.0214)</td>
</tr>
<tr>
<td>N2539</td>
<td>.0316 (.0206)</td>
<td>.0846*** (.0214)</td>
</tr>
<tr>
<td>N60</td>
<td>.0877*** (.0231)</td>
<td>.114*** (.0270)</td>
</tr>
</tbody>
</table>

Standard errors of the estimates are in parentheses below the coefficients. * indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level using a two-tailed t-test. Sample size.
obtaining maximum likelihood estimates are contained in Table 17. The
same measures for assessing the statistical reliability of the parameter
estimates and the same goodness-of-fit measures as used in the logit

Table 17. Binary logit results for the main station demand models*.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Base rate area</th>
<th>Mileage zone greater than 1B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phone vs no phone</td>
<td>Phone vs 1-party vs multi-party</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.6017</td>
<td>0.2880</td>
</tr>
<tr>
<td></td>
<td>(.3433)</td>
<td>(.3809)</td>
</tr>
<tr>
<td></td>
<td>[.0551]</td>
<td>[.0356]</td>
</tr>
<tr>
<td></td>
<td>[.0253]</td>
<td>(.0846)</td>
</tr>
<tr>
<td>1H0748ST</td>
<td>1.2099</td>
<td>.0961**</td>
</tr>
<tr>
<td></td>
<td>(.0129)</td>
<td>(.0119)</td>
</tr>
<tr>
<td></td>
<td>(.0129)</td>
<td>(.0119)</td>
</tr>
<tr>
<td></td>
<td>(.0767)</td>
<td>(.0767)</td>
</tr>
<tr>
<td></td>
<td>(.0112)</td>
<td>(.0112)</td>
</tr>
<tr>
<td></td>
<td>-.0940</td>
<td>-.1621</td>
</tr>
<tr>
<td></td>
<td>(.0345)</td>
<td>(.0345)</td>
</tr>
<tr>
<td></td>
<td>(.1116)</td>
<td>(.1116)</td>
</tr>
<tr>
<td></td>
<td>1.1780**</td>
<td>1.1780**</td>
</tr>
<tr>
<td></td>
<td>(.1451)</td>
<td>(.1451)</td>
</tr>
<tr>
<td></td>
<td>(.1398)</td>
<td>(.1398)</td>
</tr>
<tr>
<td>8</td>
<td>-.3153**</td>
<td>-.2805**</td>
</tr>
<tr>
<td></td>
<td>(.0607)</td>
<td>(.0525)</td>
</tr>
<tr>
<td></td>
<td>(.0018)</td>
<td>(.0026)</td>
</tr>
<tr>
<td></td>
<td>(.0137)</td>
<td>(.1013)</td>
</tr>
<tr>
<td></td>
<td>(.0645)</td>
<td>(.0595)</td>
</tr>
<tr>
<td></td>
<td>(.0023)</td>
<td>(.0125)</td>
</tr>
<tr>
<td></td>
<td>.1422</td>
<td>.2590</td>
</tr>
<tr>
<td></td>
<td>(.0979)</td>
<td>(.1007)</td>
</tr>
<tr>
<td></td>
<td>(.0130)</td>
<td>(.0191)</td>
</tr>
<tr>
<td></td>
<td>(.0013)</td>
<td>(.0093)</td>
</tr>
<tr>
<td></td>
<td>(.0292)</td>
<td>(.0310)</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.0110)</td>
</tr>
<tr>
<td>11517</td>
<td>.1422</td>
<td>.2590</td>
</tr>
<tr>
<td></td>
<td>(.0979)</td>
<td>(.1007)</td>
</tr>
<tr>
<td></td>
<td>(.0130)</td>
<td>(.0191)</td>
</tr>
<tr>
<td></td>
<td>(.0013)</td>
<td>(.0093)</td>
</tr>
<tr>
<td></td>
<td>(.0292)</td>
<td>(.0310)</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.0110)</td>
</tr>
<tr>
<td>11821</td>
<td>-.3189*</td>
<td>-.3509</td>
</tr>
<tr>
<td></td>
<td>(.0861)</td>
<td>(.0822)</td>
</tr>
<tr>
<td></td>
<td>-.0020</td>
<td>[.010]</td>
</tr>
<tr>
<td>12959</td>
<td>-.3009</td>
<td>-.2232</td>
</tr>
<tr>
<td></td>
<td>(.0818)</td>
<td>(.0818)</td>
</tr>
<tr>
<td></td>
<td>[.0026]</td>
<td>[.0026]</td>
</tr>
<tr>
<td>160</td>
<td>.7113**</td>
<td>.8999**</td>
</tr>
<tr>
<td></td>
<td>(.1043)</td>
<td>(.1077)</td>
</tr>
<tr>
<td></td>
<td>[.0430]</td>
<td>[.1110]</td>
</tr>
<tr>
<td>22</td>
<td>-.1973</td>
<td>.2342</td>
</tr>
<tr>
<td></td>
<td>(.1262)</td>
<td>(.1231)</td>
</tr>
<tr>
<td></td>
<td>[.0181]</td>
<td>[.0209]</td>
</tr>
<tr>
<td>1683</td>
<td>-.1973</td>
<td>.2342</td>
</tr>
<tr>
<td></td>
<td>(.1262)</td>
<td>(.1231)</td>
</tr>
<tr>
<td></td>
<td>[.0181]</td>
<td>[.0209]</td>
</tr>
</tbody>
</table>

Log likelihood -366.9567 -357.3521 -240.8666 -161.6955
Likelihood ratio index .211 .196 .171 .148
Likelihood ratio statistic 195.9396 174.416 99.276 56.0344
Percent correctly predicted 82.49 81.34 74.78 74.37

Table 17 (Continued)
Table 17 (Continued)

a Asymptotic standard errors are in parentheses below the coefficients.

b **Signifies asymptotic significance at the 10% level, *** signifies asymptotic significance at the 5% level, and **** signifies asymptotic significance at the 1% level using a two-tailed t-test.

c Effects of a unit change in each explanatory variable on the probability that the dependent variable equals unity, evaluated at the probability associated with the mean household of the sample, are in parentheses below the standard errors.

d Sample size.

The analysis of the demand for extension telephones are again employed.

For each of the models contained in Table 17, the effect of a unit change in each explanatory variable on the probability that the dependent variable equals one, evaluated at the probability associated with the mean household of the sample, is contained in brackets below the standard error. These numbers are useful for comparing the logit and ordinary least squares estimates of the magnitude of effect of the explanatory variables. As an aid in interpreting the results, the sample mean and standard deviation for each explanatory variable used in each model are shown in Table 18, and the simple correlation coefficients between explanatory variables with a correlation coefficient greater than .4 are contained in Table 19.

Before the results for each explanatory variable in each model are interpreted, the estimates obtained using the two statistical techniques are compared. The signs, magnitude and significance of the parameters of Models 3 and 4 are very similar for the ordinary least squares and binary logit estimations. Although the signs and significance of the parameters are similar, the magnitudes of the estimated coefficients in Tables 16 and 17 are different for Models 1 and 2. This result is not surprising. In the two phone versus no phone models, most

### Table 18. Sample means and standard deviations for the explanatory variables used in the main station demand models of Tables 16 and 17

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Base rate area</th>
<th>Mileage zone greater than 10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-party vs no phone</td>
<td>2-party vs no phone</td>
</tr>
<tr>
<td></td>
<td>Phone vs no phone</td>
<td>multi-party</td>
</tr>
<tr>
<td>INC74EST</td>
<td>10.6908 (8.5526)</td>
<td>10.3724 (8.7255)</td>
</tr>
<tr>
<td>PRMT1</td>
<td>7.4330 (6.6955)</td>
<td>13.5468 (2.2736)</td>
</tr>
<tr>
<td>PRMT2</td>
<td>--</td>
<td>10.0797 (5.0645)</td>
</tr>
<tr>
<td>PRMT4</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>PRMT24AV</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>M18</td>
<td>.3679 (.7442)</td>
<td>.5133 (.9111)</td>
</tr>
<tr>
<td>M924</td>
<td>.3459 (.7456)</td>
<td>.4195 (.8523)</td>
</tr>
<tr>
<td>M1527</td>
<td>.2044 (.5513)</td>
<td>.2514 (.5540)</td>
</tr>
<tr>
<td>M621</td>
<td>.2442 (.5536)</td>
<td>.3372 (.6488)</td>
</tr>
<tr>
<td>N2259</td>
<td>1.3333 (.9174)</td>
<td>1.4195 (.9064)</td>
</tr>
<tr>
<td>N60</td>
<td>.4937 (.7247)</td>
<td>.4890 (.7334)</td>
</tr>
<tr>
<td>N22</td>
<td>--</td>
<td>1.1272 (.4297)</td>
</tr>
<tr>
<td>SOMERS</td>
<td>.3646 (.3710)</td>
<td>.2167 (.4122)</td>
</tr>
<tr>
<td>COMPHS</td>
<td>.2191 (.4138)</td>
<td>.2352 (.4424)</td>
</tr>
<tr>
<td>SOMECOL</td>
<td>.1099 (.3110)</td>
<td>.0662 (.2423)</td>
</tr>
<tr>
<td>SOMECOL</td>
<td>.1447 (.3519)</td>
<td>.0646 (.2055)</td>
</tr>
<tr>
<td>COL</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>SELFEXPL</td>
<td>.0776 (.2676)</td>
<td>.0672 (.2505)</td>
</tr>
<tr>
<td>FAMPOP</td>
<td>.0283 (.1639)</td>
<td>.1379 (.3450)</td>
</tr>
</tbody>
</table>
Table 18 (Continued)

<table>
<thead>
<tr>
<th>Base rate area</th>
<th>Mileage zone greater than 18</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Phone vs. no phone</td>
</tr>
<tr>
<td>UNEMPLOY</td>
<td>.2621 (.4400)</td>
</tr>
<tr>
<td>BLACK</td>
<td>.3952 (.4981)</td>
</tr>
<tr>
<td>CSAVG</td>
<td>15.4465 (8.5772)</td>
</tr>
<tr>
<td>AVGHKMT1</td>
<td>--</td>
</tr>
<tr>
<td>AVGHKMT4</td>
<td>--</td>
</tr>
</tbody>
</table>

Standard deviations are in parentheses below the means.

of the observations lie at the extreme right-hand tail of the sigmoid-shaped response curve. In such a situation the linear probability model is clearly inappropriate and will produce biased parameter estimates. Only the binary logit results presented in Table 17 are discussed in the evaluation of the estimates for each explanatory variable.

The estimated coefficients in Table 17 can be used to predict the probability of a given household renting a particular grade of residential main station telephone service. For a given set of values for the explanatory variables for a household located within a base rate area, the estimated parameters for Model 1 together with equation (31), where PRICE = PRMT1, can be used to predict the probability that the household rents a one-party main station, the only grade of service available.

The estimated parameters for Models 2, 3, and 4 can be used to determine the probability of a household located outside mileage 18 choosing each of the available alternatives. For a given vector of values for the explanatory variables, the probabilities of a rural household selecting no phone service, a one-party line, a two-party line, and a four-party line, respectively, are given by the following expressions:

\[
Pr(\text{no phone}) = 1 - Pr(\text{PHONE}=1), \\
Pr(\text{1-party line}) = Pr(\text{PHONE}=1)Pr(\text{RMT1}=1).
\]

Table 19. Simple correlation coefficients between the explanatory variables used in the main station demand models of Tables 16 and 17

<table>
<thead>
<tr>
<th>(a) Base rate area, phone versus no phone model:</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRMT1, N60, INC74EST, UNEMPLOY, N1517</td>
</tr>
<tr>
<td>CSAVG, N60, N2259, 914, SELFEMPL, N60</td>
</tr>
<tr>
<td>(b) Mileage zone greater than 10, phone versus no phone model:</td>
</tr>
<tr>
<td>INC74EST, UNEMPLOY, N60, N914, SELFEMPL, N60</td>
</tr>
<tr>
<td>(c) Mileage zone greater than 10, one-party versus multi-party model:</td>
</tr>
<tr>
<td>AVGHKMT4, PRMT1, PRMT2, AVGHKMT2, CSAVG</td>
</tr>
<tr>
<td>(d) Mileage zone greater than 10, two-party versus four-party model:</td>
</tr>
<tr>
<td>N914, PRMT2, INC74EST, N1517, SELFEMPL, N60</td>
</tr>
</tbody>
</table>

Simple correlation coefficients are presented only for pairs of variables with a correlation coefficient of .4 or more.
Chapter 5

Pr(2-party line) = Pr(PHONE=1)[1-Pr(RMT1=1)]Pr(RMT2=1), (39)
Pr(4-party line) = Pr(PHONE=1)[1-Pr(RMT1=1)][1-Pr(RMT2=1)], (40)

where Pr(PHONE=1), Pr(RMT1=1), and Pr(RMT2=1) are defined in equations (31) with PRICE-PMT1, equation (33), and equation (35), respectively.

The binary logit estimates of the effect of each of the explanatory variables in each of the models will now be discussed. The results for the phone versus no phone models are evaluated first; then the results for the models representing choice of grade of service by rural households are examined.

Phone versus No Phone

Income. Looking first at the coefficient of INC74EST in Models 1 and 2 of Table 17, the phone versus no phone choice for urban and rural households, respectively, it can be seen that income has a highly significant effect on the probability of a household subscribing to telephone service and that the magnitude of this effect is similar for households located inside a base rate area and for households located outside mileage zone 1B. Using an expression similar to equation (29), the income elasticity of demand for telephone service, evaluated for the mean household, is calculated to be .15 for both urban and rural households. Thus, the demand for main stations, as expected, is income inelastic.

The magnitude of the income coefficient estimated for the phone versus no phone models compares favorably with empirical results obtained in two other studies. Using ordinary least squares regression analysis and 1970 Census data to estimate a linear probability model of telephone availability, Perl (1975, p. 65, see footnote 14) estimates that a $1,000 increase in income increases the probability of a household having a telephone available by 0.4. This is slightly higher than the estimated effects of a $1,000 increase in income on the probability of the mean household renting a telephone obtained here with Models 1 and 2 (.0128 and .0119, respectively). Using aggregate time-series data for Bell Canada, Waverman (1973, p. 39, see footnote 9) estimated a linear demand function for residential main telephones by regressing the number of residential main stations on the number of households, personal disposable income, and a proxy variable for the price of telephone service. The income elasticity of demand was calculated to be .15, which is the same as that estimated here. However, Waverman estimated a short-run income elasticity whereas in this study a long-run income elasticity is being estimated.

An ordinary least squares regression re-estimation of the linear probability version of Models 1 and 2 with INC74EST replaced by dummy variables representing seven different income intervals revealed, as expected, a sigmoid-shaped relationship between the probability of renting a telephone and income. This result provides additional support for employing the binary logit model for studying the phone versus no phone choice.

Price. The binary logit results of estimating Models 1 and 2 indicate that price of telephone service (PMT1 for households located within a base rate area and PMT2 for households located outside mileage zone 1B) does not have a significant effect (at the 10 percent level) on the probability of a household renting a main station. In other words, the hypothesis that the price elasticity of demand for main telephones is equal to zero cannot be rejected.

The insignificance of the price variable in the phone versus no phone models conflicts with the findings in the two studies mentioned above. Perl (1975, p. 65, see footnote 14) found price to have a
significant effect on the probability of a household having a telephone available. His ordinary least squares analysis of 1970 Census data suggests that a $1 increase in the minimum price of telephone service reduces this probability by .016. Waverman (1973, p. 39, see footnote 9) found price to be a significant factor affecting the demand for residential main telephones and estimated the price elasticity of demand to be -.12. However, it is believed that more appropriate data and statistical techniques are being used in this dissertation.

One possible explanation for the insignificance of the price variable in Models 1 and 2, Table 17, is that equation (31) is misspecified. The statistical model used here assumes that the effect of price on the probability of renting a telephone does not depend on the value of any of the other explanatory variables (i.e., it assumes the absence of any interactions). However, one would expect, and Perl (1975, p. 36, see footnote 14) has empirically verified, that the effect of price on the probability of renting a main telephone decreases as household income increases. The coefficient of the price variable in Models 1 and 2, Table 17, actually represents the average of the price effects observed at all income ranges.

The possible existence of a price-income interaction effect cannot be investigated by adding a specific interaction term—the product of income and price—to the phone versus no phone models because of the severe multicollinearity which it produces. Thus, ordinary least squares regression was used to estimate the linear probability version of Model 1 and Model 2 for subgroups of the data set consisting of households falling within each of the following four income categories: less than $6,000, $6,000-$10,000, $10,000-$15,000, and greater than or equal to $15,000. The coefficient of the price variable has the correct sign and is significantly different from zero at the 10 percent level in only one of the eight equations. Thus, in contrast to Perl's findings, a significant income-price interaction effect is not found to exist in the phone versus no phone choice, and it is assumed that this is not the cause of the insignificant price variable in Models 1 and 2.

Two alternative specifications for the price variable in the phone versus no phone model for rural households were tested to determine if the manner in which price is measured is the cause of the insignificant coefficient for this variable in Model 2. Using ordinary least squares regression, the linear probability version of Model 2 was re-estimated using the following two specifications for price: price of four-party service and the average of the price of 1-, 2-, and 4-party service. When the former specification is used, price is measured in a way similar to that of Perl (1973, see footnote 14), where the price variable is an estimate of the lowest cost of which telephone service can be attained. Neither of these alternative specifications for price proved to be significant at the 10 percent level.

Noneconomic Variables. Of the six household composition variables included in Models 1 and 2 of Table 17, only N18 and N50 have coefficients significantly different from zero at the 10 percent level for both models. The negative coefficient on N18 implies that the specific effect (i.e., need for telecommunications services) of an additional family member eight years of age or younger is outweighed by the income effect (i.e., the relative decrease in family income due to an additional member). The significant positive coefficient for the variable representing the number of family members 60 years of age or older (N50) verifies the hypothesis that the elderly have a greater need for telephone service than younger family members. In conclusion, except for persons in the youngest and oldest age categories, household size is not an important factor affecting a household's decision to rent a main telephone.
Turning now to the parameter estimates for the remaining taste variables in Models 1 and 2, Table 17, it can be seen that the probability of renting a main telephone is directly related to the educational level attained by the head of the household. Likelihood ratio tests of the equality between the coefficients of the education variables could not be performed due to the high costs of estimating binary logit probability models using maximum likelihood techniques. The probability of a black family subscribing to telephone service is significantly (at the 1 percent level) lower than for a white family, ceteris paribus. This could be the result of black families being unable to pay the advance payments and deposits due to their lower asset holdings and more limited access to credit. The results also reveal that, as hypothesized, rural households headed by a farm operator are more likely to have a telephone than rural families with heads in other occupations.

The coefficient of calling scope (CSAVG) has the theoretically anticipated sign in both Models 1 and 2 but is significantly different from zero only for Model 2. Its coefficient is not quite significant at this level in Model 1 due, perhaps, to the high degree of correlation (.51) between PRMT4 and CSAVG resulting from the “value of service” pricing principle followed by NCUC referred to previously. Irish (1974, see footnote 12) did not find calling scope to have a significant effect on the demand for one-party and two-party residential main stations within the base rate area of North Carolina exchanges.

One-Party versus Multi-Party and Two-Party versus Four-Party Service

The binary logit results of estimating the one-party versus multi-party model, presented in Model 3, Table 17, and the two-party versus four-party model, shown in Model 4, Table 17, will now be examined. The first observation to make is that, although income is an important determinant affecting a rural household’s decision to receive telephone service, it is not a significant factor affecting choice of grade of service. The coefficient of INC74EST is not significantly different from zero at the 10 percent level in either Model 3 or Model 4.

Price, which does not exert a significant effect on a household’s decision to rent a main telephone, is an important factor affecting a rural household’s selection of grade of service. The coefficients of PRMT2 and PRMT2AV in Model 3 and PRMT3 and PRMT4 in Model 4 have the theoretically anticipated sign and are highly significant. The results indicate that, for the mean household located outside mileage zone 18 that has elected to receive telephone service, a $1 increase in the monthly rental rate for individual line service (PRMT2), ceteris paribus, decreases the probability of selecting this grade of service by .16. For rural households choosing multi-party service, a $1 increase in the price of four-party service (PRMT4), ceteris paribus, increases the probability of the mean household choosing a two-party line by .16.

Turning now to the estimated parameters for the noneconomic variables, Models 3 and 4 in Table 17 indicate that, although the coefficients of some of the household composition variables are significant, in general the effect of household size and composition on choice of grade of service is erratic and un systematic. The probability of selecting a higher quality grade of service increases with the amount of schooling attained by the head. Occupation of the head has little effect on choice of grade, with the exception that families with heads earning no income during 1974 are much less likely to have a one-party line than households with an employed head. Although blacks are less likely to have a telephone than whites, those that do are more likely than whites to select a higher quality grade of service.
The binary logit results for Models 3 and 4 indicate that the average number of hookups on two-party and four-party lines for the exchange in which the household is located does not exert a significant (at the 10 percent level) effect on the choice of grade of service. The insignificance of these variables may be due to the small amount of variation existing for AVGHKMT2 and AVGHKMT4 in the sample used in these calculations—an indication that the telephone company attempts to keep its multi-party lines full.

**Goodness of Fit**

The goodness-of-fit measures provided with the binary logit results of Table 17 are quite satisfactory and indicate that the four main station demand models fit the data better than the extension demand models of Table 10. The likelihood ratio statistics reveal that the null hypothesis that the coefficients of all explanatory variables except the constant are equal to zero can be rejected at the 1 percent level for each model. The likelihood ratio index and the percent correctly predicted show that the phone versus no phone models are providing better fits to the data than the choice of grade of service models.

**Summary and Conclusions**

The objective of this study was to use data on individual households to determine which factors affect the demand for several types of residential telephone service and to estimate the magnitude of these effects. The explanatory variables considered can be grouped into three categories—economic factors, such as household income and prices; traditional noneconomic factors, such as household size and composition, degree of urbanization, race, and social class; and noneconomic factors specific to the services under analysis, such as geographical dispersion of family, size of housing unit, and calling scope. In the first stage of this analysis the factors affecting a household's expenditures on long-distance calls were analyzed. A double-logarithmic demand function was formulated and estimated using ordinary least squares regression.
Chapter 8

The factors affecting a residential telephone subscriber's decision to rent an extension station were examined in the second phase of the study. Due to the (0,1) nature of the dependent variable in this analysis, ordinary least squares regression would produce heteroscedasticity and is not the appropriate statistical technique to use in estimating the model developed to explain this choice. Thus, binary logistic analysis, which hypothesizes a sigmoid-shaped relationship between the probability of renting an extension telephone and a linear combination of the explanatory variables, was used in the empirical analysis.

In the third phase of this dissertation the factors affecting a household's decision to subscribe to telephone service and its choice among available grades of service (e.g., 1-party, 2-party, or 4-party) were investigated. Since only individual line service is available to households located within the base rate area of an exchange, a binary logistic probability model was developed and estimated to explain the choice between not subscribing to telephone service and renting a one-party main station. Three grades of service are available to rural households living outside mileage zone 1B. In deciding whether or not to subscribe to telephone service and in selecting grade of service, these households are assumed to follow a sequential decision-making process. In the first stage, the household decides whether or not to rent a main station. Households that have elected to receive telephone service next decide whether to rent a one-party line or a multi-party line. Households selecting multi-party service then, in stage three, choose either two-party or four-party service. Due to the dichotomous nature of the dependent variable in each of these models, a series of binary logistic probability models was formulated and estimated.

The data set analyzed in this study was created by merging information gathered in interviews with a random stratified sample of 2,113 households in eastern North Carolina with local and long-distance telephone service consumption data provided by the telephone company serving the area. Additional variables representing telephone system attributes were constructed from information on file with the North Carolina Utilities Commission. Data on various subgroups of the households were used to estimate each of the above discussed demand functions.

The results obtained in this study are encouraging. Many variables were found to be significant and with the correct sign. The models fitted the data well considering that an individual household was the unit of observation. The results obtained are summarized in the first section of this chapter. The implications of the findings for regulation and planning by telephone companies are discussed in the following section. Some suggestions for future research conclude the chapter.

Summary

Although the empirical findings of this study are based on data for only a small section of the population, they nevertheless provide insight into the effects of many factors on the demand for various types of residential telephone service. The results of this study are summarized by describing the effect of each explanatory variable on the demand for each type of residential telephone service studied.

Economic Variables

Utility theory suggests that income and prices are important determinants of the amount a household consumes of some service. Although total family gross income was found to have a significant and positive effect on the probability of a family subscribing to telephone service, it is not a significant factor affecting the choice of grade of service by households living outside a base rate area. The income elasticity of demand for telephone service is calculated to be .15.
which is very similar to the estimates obtained by previous researchers. As expected, households with a higher level of income, *ceteris paribus*, are more likely to rent an extension telephone. They also spend more on long-distance calls. The income elasticities of demand for extension telephones and toll calls are estimated to be .21 and .17, respectively. The income elasticity of demand for extension telephones corresponds with the findings of other investigators, but the income elasticity of demand for toll calls is much smaller than the best available estimates. The low income elasticity of demand for toll calls estimated here is not believed to be due solely to using current rather than permanent income as the measure of purchasing power.

With the data analyzed here cross-sectional price variation exists only for residential main station monthly rental rates. Price of telephone service was found not to have a significant effect on a household's decision to rent a main station. In addition, the effect of price was found not to be dependent on the level of income (*i.e.* there is no price-income interaction effect). Both of these results conflict with the findings of previous researchers using less appropriate data and statistical techniques. Rates for the available grades of service were found, however, to be significant determinants of the choice of grade of service. For the mean rural household that has elected to receive telephone service, a $1 increase in the monthly rental rate for individual line service, *ceteris paribus*, decreases the probability of selecting this grade of service by .16. For rural households choosing multi-party service, a $1 increase in the price of four-party service, *ceteris paribus*, increases the probability of the mean household choosing a two-party line by .16.

**Traditional Non-economic Variables**

In addition to economic variables, each demand function formulated in this analysis contains several variables frequently used in empirical demand studies to take account of differences in tastes and preferences among consumers. The size of a household and the age-sex composition of its members are usually considered to be the most important non-economic variable influencing the demand for a good or service. Of the six variables used to measure the size and composition of a household, only two were found to have a significant effect on the decision to rent a main station. Additional family members eight years of age or younger reduce the probability of receiving telephone service, whereas an additional family member 60 years of age or older increases this probability. Household size and composition exert little systematic effect on a household's selection of grade of service. Household members 22 years of age or older increase the probability of a family renting an extension telephone by .12, and an additional person 16 years of age or older increases household expenditures on toll calls by 25 percent, all other things remaining the same.

Tastes and preferences of consumers may also depend on the social class to which the family belongs. Education and occupation of the head of the household were used as proxies for social class in this study. It was discovered that the more educated the head, the more likely the household is to subscribe to telephone service, select a higher quality grade of service, rent an extension telephone, and spend more on long-distance calls. It is possible that this direct relationship between education and the demand for telephone services may not be reflecting a difference in taste for telephone services by families in different social classes, but rather a difference in permanent income. While education has a significant effect on the demand for all types of residential telephone service, occupation of the family head, in general, does not influence the demand for any type of service.
Differences in the consumption of telephone services among the households included in the sample may also be attributed to differences in their degree of urbanization and to race. The effect of urbanization on a household’s decision to rent a main station and its selection of grade of service could not be analyzed since almost all households located within a base rate area are classified as urban households and most all families living outside mileage zone 1B are classified as rural families. The probability of a rural household renting an extension telephone was found to be .15 less than that for a similarly situated urban family. Rural families also spend 22 percent less on toll calls than urban families, ceteris paribus. Race was found to be a strong and significant determinant of the demand for each type of telephone service analyzed. While black families have a lower probability of renting a telephone, those that do are more likely to select a higher quality grade of service and to rent an extension telephone than white families, all other things being equal. Blacks also spend 56 percent more on long-distance calling than whites, ceteris paribus.

Special Noneconomic Variables

In addition to the usual noneconomic variables, several special explanatory variables were included in the demand models to represent telephone system attributes and household characteristics which are hypothesized to affect the consumption of the particular services under analysis. One such variable, recency of residency, was not found to have a significant effect on toll expenditures. Toll calling was found, however, to increase as the geographical dispersion of immediate family relatives living outside the household’s local calling area increases. The empirical results also reveal that families using the telephone more intensively for local communication also make more

Summary and Conclusions

As expected, families living in larger housing units have a greater probability of renting an extension station, ceteris paribus.

Although the evidence is not conclusive, the empirical results of this research provide some indication that the probability of a household subscribing to telephone service is directly related to the size of the calling scope of the exchange in which it is located. Calling scope was found to have a significant and inverse relationship with toll calling. A rural subscriber’s choice of grade of telephone service was not found to be dependent on the average number of hookups on two-party and four-party lines for the exchange in which it is located.

Policy Implications

Although the main purpose of this study was to estimate the parameters contained in the demand functions for several types of residential telephone services, the findings have several important implications for policy decisions made by regulatory authorities and telephone company planners.

Implications for Regulation

The most useful discovery in this research for state utility regulators is the finding that the price elasticity of demand for telephone service is not significantly different from zero. This result implies that changes in telephone rates, at least within the range of rates observed in the data analyzed in this study, will not significantly change the number of telephone subscribers. The results of this study do reveal, however, that the rate structure for main station telephone service facing rural subscribers is an important factor affecting their choice of grade of service. Changes in the rates facing these subscribers, due either to changes in base rates or mileage zone charges,
may result in subscribers upgrading or downgrading their class of service.

In rate cases it is usually assumed that changes in main station rates will not change the number of subscribers and that existing subscribers will not change their grade of service. The first of these two hypotheses cannot be rejected with the data analyzed in this study. The second one, however, can be. This implies that predicting changes in grades of service which may occur due to rate changes will result in more accurate forecasts of the effects of rate changes on telephone company revenue and, hence, improvements in regulation.

Although the effects of rate changes on selection of grade of service may be slight and practicality may dictate ignoring such changes in rate-making proceedings, the procedure to be followed in assessing the effects of rate changes will be described, mainly to demonstrate the manner in which predictions can be made using the results of estimating sequential binary logit models. Suppose a state utility commission wishes to predict the effect of a change in the monthly base rate for two-party residence telephone service, \( \text{ceteris paribus} \), on the number of one-party, two-party, and four-party residential main stations rented by households located outside the base rate area of a particular telephone company. From equations (38), (39), and (40), the initial probability of a rural residential telephone subscriber renting a one-party, two-party, and four-party main station, respectively, can be expressed as follows:

\[
\begin{align*}
\Pr(\text{one-party line}) &= \Pr(RMT1=1), \\
\Pr(\text{two-party line}) &= (1-\Pr(RMT1=1))\Pr(RMT2=1), \\
\Pr(\text{four-party line}) &= (1-\Pr(RMT1=1))(1-\Pr(RMT2=1)),
\end{align*}
\]

where \( \Pr(RMT1=1) \) and \( \Pr(RMT2=1) \) are as defined in equations (33) and (35), respectively. The effects of a small change in the price of two-party service on the probability that a rural telephone subscriber will select one-party, two-party, and four-party service, respectively, can be expressed as follows:

\[
\begin{align*}
\Delta \Pr(\text{one-party line}) &= \gamma_0 \Pr(RMT1=1)\Delta \Pr(RMT1=0), \\
\Delta \Pr(\text{two-party line}) &= \gamma_0 \Pr(RMT1=1)\Pr(RMT1=0)\Pr(RMT2=1) + \\
&\quad + \delta \Pr(RMT2=1)\Pr(RMT2=0)(1-\Pr(RMT1=1)), \\
\Delta \Pr(\text{four-party line}) &= -\gamma_0 \Pr(RMT1=1)\Pr(RMT1=0)(1-\Pr(RMT2=1)) \\
&\quad - \delta \Pr(RMT2=1)\Pr(RMT2=0)(1-\Pr(RMT1=1)).
\end{align*}
\]

where \( \gamma_0 \) and \( \delta \) are as defined in equations (34) and (36), respectively.

For a given vector of values for the explanatory variables for a particular household, the estimated parameters presented in Models 3 and 4 of Table 17 and equations (44), (45), and (46) can be used to predict the effect of a small change in the price of two-party service on the probability of a particular household selecting each of the available grades of service. Regulatory authorities, however, are not interested in the reaction of a particular individual to rate changes. They want to predict the change in the total number of one-, two-, and four-party main stations rented by an aggregate of individuals, such as those located within a particular exchange or service area of a particular telephone company. As equations (44), (45), and (46) indicate, the predicted changes in the probability of selecting each available grade of service are not the same for all individuals but depend on the subscriber's initial probabilities of selection. Thus, to predict the expected change in the aggregate number of subscribers selecting each grade of service, one cannot simply extrapolate the estimated changes in probabilities for a
representative subscriber, such as the one having the initial probabilities corresponding to the mean household of the sample. Instead, the change in probability corresponding to each initial probability must be weighted by the relative frequency distribution of initial probabilities across the population whose behavior is to be predicted. Calculating the average change in the probability of selecting each grade of service in this manner also takes account of difficulties encountered in using logit results calibrated on one population to make predictions about another population.

Several predictive techniques which incorporate the initial probabilities of choice are available. Westin (1974) describes a procedure for calculating the relative frequency distribution of initial probabilities from assumptions concerning the distribution of explanatory variables in the population. Alternatively, one could just choose a convenient functional form for the frequency distribution. Using a method suggested by Westin such a relative frequency distribution could be combined with parameter estimates, such as those contained in Models 3 and 4 of Table 17, to estimate the expected proportion of households in the population that will select each grade of service.

Another method for predicting the effects of a change in the price of two-party service, which also takes account of the initial probabilities of selection, is to follow Demenich and McFadden (1975, p. 82) and use summary statistics for the population under analysis and their distribution to generate a hypothetical random sample. With data on the characteristics of each hypothetical household and price for each grade of service available, the estimated coefficients of Models 3 and 4 in Table 17 and equations (41), (42), and (43) could be used to calculate the probability of each household selecting each grade of service. These probabilities can then be averaged to determine the expected proportion of the population selecting each grade. Multiplying these proportions by the population of the service area of the company produces the expected demand for 1-party, 2-party, and 4-party residential main stations. Alternatively, if the sample used to calibrate the models is random, these data and the procedure described above may be used to construct an estimate of aggregate demand. To determine if the data set used in the logit analyses of this study may be used for forecasting, the randomness of each sample was evaluated. The subgroup of households used to calibrate each model was found to be representative of the eight-county area surveyed.

The second approach described above would also be the suggested technique for predicting the effects on choice of grade of service of a simultaneous change in the rates for all three grades. Usually, when a state utilities commission changes rates, the rates for all grades of service, not just one, are changed. In such a situation, data on the characteristics and price of each available grade of service for a random sample of households from the population under analysis could be combined with the parameter estimates contained in Models 3 and 4 of Table 17 and equations (41), (42), and (43) to calculate the probability of each household selecting each grade of service available. Multiplying the size of the population by the average probability of selecting 1-, 2-, and 4-party service, calculated for the sample, produces the expected aggregate demand for each grade.

The results of this study may also be useful for studying other regulatory policy changes. With the data gathered in this research, a model could be developed and estimated to determine if expansion of a base rate area (i.e., elimination of two- and four-party main station service in areas currently outside a base rate area) would reduce the number of residential telephone subscribers. Casual empiricism suggests
that it will not. The empirical findings in this research could also be used to predict the increase in the number of telephone subscribers and the decrease in toll revenue that would result from an increase in the size of a calling scope, due, perhaps, to implementation of Extended Area Service.

The results of this econometric study also have some interesting implications for the adoption of "lifeline" or "usage sensitive" pricing of residential main station telephone service. Instead of paying a monthly flat rate which permits a subscriber to make an unlimited number of local calls, with this alternative pricing scheme a telephone subscriber pays a lower monthly rate but is charged for each additional call made after some minimum number. The purpose of such a pricing policy is to increase subscribership among low income families and the elderly who must have connection with the outside world. This pricing scheme is already in use in a few areas of the country and is being considered for adoption in many others.

The results for Models 1 and 2 in Table 17 indicate that low income families are less likely to have a telephone and that older individuals do derive more utility from telephone service than younger persons. The results also suggest, however, that lowering the price of telephone service will not attract new subscribers, not even among low income households. Although these findings pertain to a flat rate pricing system, they do at least suggest that adoption of lifeline rates may not achieve the goal of increasing the number of subscribers. Perhaps the major factor inhibiting low income households from renting a telephone is not the monthly rental rate, but rather the deposits and advance payments required from applicants for service with poor credit ratings.

This study's results would also be of use in formulating and/or calibrating models designed to analyze the effects on usage of converting from a flat rate to a two-part tariff pricing system. Mitchell (1976) has formulated such a model and calibrated it using Perl's (1975, see footnote 14) estimates of the effect of telephone rates on telephone availability. The results of this study could also be used to calibrate such a model.

The empirical results obtained in this study also have some interesting implications for the pricing of long-distance telephone service by state and federal regulatory authorities. In the past these agencies have permitted long-distance rates to exceed costs in order to subsidize local telephone service. If this subsidy is based on the notion that long-distance calling is a luxury item whereas local service is more of a necessity, the income elasticities calculated in this study suggest that this assumption is unfounded. If goods and services with income elasticities less than one can be classified as necessities, an income elasticity of .17 for toll service and .15 for main telephone, as estimated in this study, suggest that both of these services are necessities. Both of these income elasticities are lower than those usually found for generally regarded necessities such as food, clothing, and shelter.

Implications for Telephone Company Planning

Telephone company planners may find the techniques and results of this study useful for forecasting future demand for residential telephone services. Due to the large number of noneconomic variables found to be significant determinants of the demand for residential telephone services, incorporating these variables in forecasting models should produce more satisfactory predictions than can be obtained with the largely extrapolative techniques usually employed. Improvements in predicting future demand and, hence, future capital stock requirements would improve scheduling of investment activities so that equipment is in place as it is needed.
Since household size and composition, race, occupation, education, location, and other noneconomic factors change only gradually over time, in making short-run forecasts planners are most interested in predicting the effects of changes in income and prices. Although the income and price elasticities calculated in this study can be used in making these forecasts, in the case of residential main stations and extensions more accurate predictions can be obtained by using one of the procedures discussed above which takes account of the initial frequency distribution of probabilities of choice for the population under consideration.

The low income elasticities of demand discovered in this research have important implications for the growth of the telephone industry. Income elasticities of .15, .21, and .17 for main stations, extension stations, and toll calls, respectively, imply that residential telephone services are necessities and that the demand for these services will not increase much as income increases over time.

The estimated effects of the noneconomic variables obtained in this research can similarly be used to predict the long-run effect on the demand for residential telephone services of such changes as increased urbanization, increased educational levels, increased geographical dispersion of families, and increases in the size of housing units. For example, future projections for population, average income per household, and values for all the other explanatory variables in the toll demand function can be combined with the estimated parameters contained in Table 3 to predict future expenditures on long-distance calls assuming the relative price of this service does not change.

**Suggestions for Further Research**

Although a number of interesting and useful results have been obtained in this study, the unique body of data assembled as part of this research remains rich in possibilities for further analysis. Some suggestions for future research will now be made.

In the future it is hoped that the 1974 toll billing statements can be coded and placed on computer tape. These additional toll data could be used to lengthen the period of observation on long-distance calling behavior or to produce a dependent variable coinciding to the same time period as the income variable. This added information would also make possible an investigation of the dynamic aspects of the demand for long-distance calls.

Another avenue for further research is to utilize the great amount of detail contained in the long-distance information provided by the telephone company to explore the factors affecting a residential toll caller's choice of time of day, class of service (person-to-person or station-to-station), method of payment (sent paid, collect, third person), etc. With the detailed information on each long-distance call that is available, the dependent variable in the toll demand function could also be converted to message-minute-miles.

The income elasticities of demand for toll calls and extension telephones calculated here could, using restricted least squares regression, be integrated with time-series data to yield estimates of the price elasticities of demand for these services.

Due to the high costs of performing logit analysis, the decision to rent a telephone and the choice of grade of main station service could not be analyzed in this study as thoroughly as the demand for the other residential telephone services. It would be interesting to estimate a polychotomous logit model of choice of grade of service and compare the results with the estimates obtained here with the sequential binary choice models. Also, it is hoped that logit analysis can be employed to search for interactions between the explanatory variables.
166 Chapter 6

included in the main station and extension demand models and to investigate the effects of expressing the explanatory variables as transformations of the raw data. It would also be interesting to test various restrictions on the independent variables included in these two models. In addition, the author plans to use the data base to conduct a simulation study to determine the effect of a simultaneous change in the rates for all three grades of service on the demand for each grade.

The detailed local service information provided by the telephone company could be used, mainly for marketing purposes, to study the factors affecting the demand for touch-call, various styles of telephone instruments, extra bells and gongs, and other pieces of miscellaneous equipment. Interesting results might also be obtained from estimating a demand function where the dependent variable is total monthly expenditures on residential telephone services (i.e., the sum of monthly local service charges and average monthly expenditures on toll calls).

References


Appendix A. Data Description

The data set analyzed in this study was created by supplementing the household interview information gathered in the Second Survey of Labor Supply in Eastern North Carolina with telephone service consumption data provided by Carolina Telephone and Telegraph Company. Additional variables were constructed from data filed with the North Carolina Utilities Commission. Each of these three data sources will now be described.

Survey Data

From mid-March 1975 to early June 1975, the Department of Economics and Business of North Carolina State University at Raleigh obtained completed interviews with 2,113 households living in the eight eastern North Carolina counties of Beaufort, Bertie, Hertford, Martin, Greene, Pitt, Lenoir, and Jones. The purpose of conducting these personal interviews was to obtain information useful for predicting the effects that mechanization of tobacco production and harvesting would have on employment and labor income in a tobacco producing region. Many of these same families were also interviewed in spring 1973, as part of the same study.

Survey specialists at the Research Triangle Institute, Research Triangle Park, North Carolina, designed and selected a random stratified sample representative of the noninstitutional housing population of this eight-county area. This was accomplished by dividing the eight-county study area into a number of carefully defined geographic strata and
then sampling each stratum by multi-stage area probability sampling methods.

The stratification scheme was a function of two factors: the region in which the household was located and its degree of urbanization (e.g., rural or urban). There were 67 strata, each containing the same number of households. After the strata were designated, six segments (small areas of land bounded by roads, creeks, rivers, or railroad tracks in rural areas and streets in urban areas) were chosen from each stratum in Green, Lenoir, and Jones counties and seven segments were selected from each stratum in Beaufort, Bertie, Hertford, Martin, and Pitt counties. In each stratum all segments had the same probability of being selected. A list was prepared of all dwelling units located within each selected segment. Systematic random sampling was used to select from each segment roster an approximately equal number of households to be interviewed. This stratified random sampling scheme provided assurance of geographic dispersion of households within the survey area beyond that which might have resulted from simple random sampling.

Attempts were made to interview over 2,817 households. Interviewers were unable to contact the occupants of 116 housing units; 331 housing units were found vacant; 119 households refused to be interviewed; and interviews could not be obtained from 148 households for various other reasons. Thus, completed interviews were obtained from 2,113 households, approximately 85 percent of the occupied housing units. This is a fairly high response rate considering that the Survey Research Center, which conducts many surveys, experiences nonresponse rates in the range of 10 percent to 25 percent.

The average length of these personal interviews was 33 minutes.

Appendix

Some of the information obtained from each household during this time is as follows:

1. Residence status (whether or not this household was also interviewed in 1973),
2. Degree of urbanization (rural or urban),
3. Race (white, black, or other),
4. Age, sex, marital status, relationship to head of household, and highest grade completed in school for each family member present in 1974,
5. Occupation, industry of employment, number of weeks worked, and pay per period for each job held during 1974 for each family member 12 years of age or older,
6. Past migration history of family (whether or not the family moved into the country after January 1, 1970),
7. 1974 family income from sources other than farming, wages and salaries, and self-employment (e.g., social security, unemployment insurance, welfare, retirement payments, etc.), and
8. Interval estimate for 1974 total family gross income (under $2,000, $2,000-3,999, $4,000-5,999, $6,000-9,999, $10,000-14,999, $15,000-19,999, $20,000-49,999, $50,000 and over, not ascertained).

In addition to the above data on the general characteristics of household members, the interviewers gathered information especially for use in studying the demand for residential telephone service. As the interviewers approached the housing units, they were to note the number of levels of living quarters (1, 2, or 3). At the end of the interview each respondent was asked if there was a telephone installed in the housing unit and, if so, they were asked to provide the telephone number. Those families subscribing to telephone service were then asked to list...
the city and state in which each member of the immediate family of the head and spouse lived, if they lived outside the household's local calling area. The relationship of each of these relatives (i.e., father, mother, daughter, son, brother, or sister of the head or of the spouse) was also recorded. The interviewees were then asked to give their written permission for CT&T to release their local and long-distance telephone records.

Using information obtained during the interview on the work effort of each family member, a continuous variable estimate of total family 1974 gross income was calculated. If any of the necessary partial information on hours worked or pay per period was missing for a household member, that variable was assigned a value equal to the mean for similar persons in the survey.

For those households that were also interviewed in spring 1973, information on estimated 1972 total family gross income and the number of each type of room in the housing unit was obtained from the First Survey of Labor Supply in Eastern North Carolina, which was also conducted by the Department of Economics and Business of North Carolina State University.

Data Provided by CT&T

Carolina Telephone and Telegraph Company, the second largest telephone company in North Carolina, provided telephone service to 97.15 percent of the households interviewed that were receiving telephone service. Of the remaining households, seven were receiving their telephone service from Norfolk and Carolina Telephone and Telegraph Company, and 40 were having their telephone service provided by the Tri-County Telephone Membership Corporation, a nonregulated telephone company.

CT&T supplied local service and toll service information for their residential subscribers that were surveyed and granted permission for the telephone company to do so. Approximately 71 percent of the 1,601 CT&T subscribers interviewed gave written permission for the telephone company to release their records. No attempt was made to obtain telephone records for the 47 households served by the other two telephone companies.

For each subscriber signing the permit slip, the following local service information was obtained:128

1. name and address to which the telephone bill is sent,
2. customer type (residence or business),
3. grade of main station service (1-, 2-, or 4-party),
4. number of inside and outside extension stations,
5. exchange and mileage zone in which the subscriber is located,
6. total monthly local service charges,
7. amount of concession for telephone company employees,
8. for each telephone set, information on the type of set (desk, compact, dial-in-handset, Ericofon, decorator, panel, etc.), color, where mounted (hand or wall), type of cord (spring or non-spring) and length, whether dial is illuminated or not, whether or not it is equipped with touch-call, etc.,
9. number of additional listings,
10. whether the telephone number is nonpublished or nonlisted, and
11. complete list of all other miscellaneous equipment (such as extra bells or gongs, jacks, automatic dialing devices, outdoor telephone sets, etc.) installed on the subscriber's premises.

The long-distance information for the CT&T subscribers surveyed was provided in three different forms. The computer tape containing the above described local service information also contained toll charges for each of the previous three billing periods for each subscriber. Since
not all subscribers have the same billing period, the calendar period to which this information corresponds is not the same for each household. These data cover a time period from February 18, 1975, through June 16, 1975.

In addition to this three-month's worth of toll data, CT&T created and supplied a computer tape containing the following information for each toll call placed by each subscriber in the sample from July 1, 1975, through November 31, 1975:

1. charge for the call, excluding federal excise tax,
2. class of toll message service (station-to-station or person-to-person),
3. type of call (sent paid, collect, credit card, or third number),
4. whether the call was operator handled or dialed direct,
5. date of call,
6. time of connection,
7. length of conversation, and
8. place and number called, etc.

The third source of long-distance information, toll billing statements, was obtained from CT&T records stored in five billing offices in eastern North Carolina. Toll statements contain the same information as listed above but are in paper form. For each subscriber in the sample, toll statements were obtained for nine or 10 billing periods in 1974. This information has not yet been coded and transferred to a computer tape.

Description of Procedures Employed in Constructing Additional Variables

Data obtained from the NCUC and the FCC were used to construct additional variables. The calculation of each of these variables will now be described.
Average Number of Hookups on Two-Party and Four-Party Lines. The average number of hookups on two-party lines for a particular exchange is the ratio of the number of two-party main stations in the exchange (including business as well as residence) to the number of two-party lines in the exchange. The average number of hookups on four-party lines has an analogous interpretation. Both of these variables were calculated from data contained in CT&T's May 31, 1975, station development report filed with the NCUC.

Calling Scope. Calling scope is the number of main stations, private branch exchange trunks (i.e., switchboards) and other equivalents that a subscriber may call toll-free. Due to Extended Area Service calling scope may not just be the number of telephones located in the subscriber's own exchange but also includes the number of telephones in the other exchanges which comprise the subscriber's local calling area. CT&T's Local Exchange Tariffs (1973, see footnote 130) were first used to determine the exchanges included in the local calling area of each exchange in the survey area. The calling scope of each exchange was then calculated from monthly station development reports filed by CT&T with NCUC. These reports show the number of main stations and private branch exchange trunks in each exchange at the end of each month. The calling scope figure used in the main station demand analysis was the average size of the calling scope for the four months of February, March, April, and May 1975. The average size of the calling scope for the 10 months from February through November 1975 was the calling scope figure used in the toll demand analysis.

Distance to Relatives. The exchange in which each subscriber was located was also used to assign each household a vertical (V) and a horizontal (H) coordinate using FCC Tariff No. 264 filed by AT&T on February 28, 1975. This tariff provides a V-H coordinate for each exchange in the United States. These coordinates are computed from the latitude and longitude of each rate center adjusted to reflect the curvature of the earth's surface. This tariff was also used to assign to each relative each household reported as living outside its local calling area the V-H coordinates corresponding to the nearest rate center. The Pythagorean Theorem was then applied, following original page 9.00 of this tariff, to compute the airline distance to each relative. This is the same procedure used to calculate the length-of-haul for a long-distance telephone call.
### Appendix B. Comparison between 1972 and 1974 Gross Family Income for Households Used in the Toll Demand Analysis

**Appendix Table 1. Comparison between 1972 and 1974 estimated income**

<table>
<thead>
<tr>
<th>1972 Estimated Income</th>
<th>1974 Estimated Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$2,000</td>
<td>3,999</td>
</tr>
<tr>
<td>2,000-3,999</td>
<td>5,999</td>
</tr>
<tr>
<td>4,000-5,999</td>
<td>9,999</td>
</tr>
<tr>
<td>6,000-9,999</td>
<td>14,999</td>
</tr>
<tr>
<td>10,000-14,999</td>
<td>29,999</td>
</tr>
<tr>
<td>15,000-29,999</td>
<td>49,999</td>
</tr>
<tr>
<td>30,000-</td>
<td>50,000</td>
</tr>
</tbody>
</table>

| <$2,000               | 2,000-4,000           |
| 2,000-4,000           | 6,000-10,000          |
| 4,000-6,000           | 10,000-15,000         |
| 6,000-9,000           | 15,000-20,000         |
| 8,000-11,000          | 20,000-30,000         |
| 10,000-14,000         | 30,000-50,000         |

Table entries are the number of observations falling into each cell.

**Notes:**
- Only those observations used in the toll demand analysis for which both 1972 and 1974 estimated income are available are included in the table.
### Appendix Table 2. Comparison between 1972 and 1974 reported income\(^a\),\(^b\)

<table>
<thead>
<tr>
<th>1972 reported income</th>
<th>1974 reported income</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2,000</td>
<td>&lt;$2,000</td>
</tr>
<tr>
<td>2,000-3,999</td>
<td>2,000-3,999</td>
</tr>
<tr>
<td>4,000-5,999</td>
<td>4,000-5,999</td>
</tr>
<tr>
<td>6,000-9,999</td>
<td>6,000-9,999</td>
</tr>
<tr>
<td>10,000-14,999</td>
<td>10,000-14,999</td>
</tr>
<tr>
<td>15,000-19,999</td>
<td>15,000-19,999</td>
</tr>
<tr>
<td>20,000-24,999</td>
<td>20,000-24,999</td>
</tr>
<tr>
<td>25,000-29,999</td>
<td>25,000-29,999</td>
</tr>
<tr>
<td>30,000-34,999</td>
<td>30,000-34,999</td>
</tr>
<tr>
<td>≥35,000</td>
<td>≥35,000</td>
</tr>
</tbody>
</table>

\(^a\)Table entries are the number of observations falling into each cell.

\(^b\)Only those observations used in the toll demand analysis for which both 1972 and 1974 reported income are available are included in the table.

### Appendix C. Randomness Analysis of the Toll and Extension Samples

### Appendix Table 3. Randomness analysis of the toll and extension samples

<table>
<thead>
<tr>
<th>Variable</th>
<th>Toll Data</th>
<th>Extension Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported 1974 gross income:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$2,000</td>
<td>51</td>
<td>44</td>
</tr>
<tr>
<td>2,000-3,999</td>
<td>98</td>
<td>104</td>
</tr>
<tr>
<td>4,000-5,999</td>
<td>111</td>
<td>123</td>
</tr>
<tr>
<td>6,000-7,999</td>
<td>188</td>
<td>228</td>
</tr>
<tr>
<td>10,000-14,999</td>
<td>201</td>
<td>220</td>
</tr>
<tr>
<td>15,000-29,999</td>
<td>188</td>
<td>197</td>
</tr>
<tr>
<td>≥30,000</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>68</td>
<td>73</td>
</tr>
<tr>
<td>Estimated 1974 gross income:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$2,000</td>
<td>41</td>
<td>30</td>
</tr>
<tr>
<td>2,000-3,999</td>
<td>83</td>
<td>80</td>
</tr>
<tr>
<td>4,000-5,999</td>
<td>97</td>
<td>94</td>
</tr>
<tr>
<td>6,000-7,999</td>
<td>111</td>
<td>111</td>
</tr>
<tr>
<td>10,000-14,999</td>
<td>226</td>
<td>247</td>
</tr>
<tr>
<td>15,000-29,999</td>
<td>235</td>
<td>246</td>
</tr>
<tr>
<td>≥30,000</td>
<td>29</td>
<td>33</td>
</tr>
<tr>
<td>Degree of urbanization:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural nonfarm</td>
<td>494</td>
<td>387</td>
</tr>
<tr>
<td>Rural farm</td>
<td>107</td>
<td>81</td>
</tr>
<tr>
<td>Urban</td>
<td>332</td>
<td>185</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Race:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>674</td>
<td>450</td>
</tr>
<tr>
<td>Black</td>
<td>260</td>
<td>231</td>
</tr>
<tr>
<td>No. family members 18 years of age or older:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>129</td>
<td>121</td>
</tr>
<tr>
<td>2</td>
<td>357</td>
<td>388</td>
</tr>
<tr>
<td>3</td>
<td>178</td>
<td>137</td>
</tr>
<tr>
<td>4</td>
<td>55</td>
<td>46</td>
</tr>
<tr>
<td>5 or more</td>
<td>19</td>
<td>9</td>
</tr>
</tbody>
</table>

\(^132\)
## Appendix Table 3 (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Toll Data Set 2</th>
<th>Data Extension Data Set 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. family members 17 years of age or younger:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>492</td>
<td>529</td>
</tr>
<tr>
<td>1</td>
<td>185</td>
<td>204</td>
</tr>
<tr>
<td>2</td>
<td>140</td>
<td>156</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>4 or more</td>
<td>47</td>
<td>49</td>
</tr>
<tr>
<td>Education of head:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade school</td>
<td>349</td>
<td>374</td>
</tr>
<tr>
<td>Some high school</td>
<td>119</td>
<td>169</td>
</tr>
<tr>
<td>Completed high school</td>
<td>220</td>
<td>243</td>
</tr>
<tr>
<td>Some college</td>
<td>100</td>
<td>111</td>
</tr>
<tr>
<td>Completed college</td>
<td>106</td>
<td>121</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Age of head:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>44</td>
<td>34</td>
</tr>
<tr>
<td>25-34</td>
<td>140</td>
<td>162</td>
</tr>
<tr>
<td>35-44</td>
<td>165</td>
<td>177</td>
</tr>
<tr>
<td>45-54</td>
<td>208</td>
<td>214</td>
</tr>
<tr>
<td>55-64</td>
<td>177</td>
<td>190</td>
</tr>
<tr>
<td>65-74</td>
<td>167</td>
<td>156</td>
</tr>
<tr>
<td>&gt;75</td>
<td>53</td>
<td>61</td>
</tr>
<tr>
<td>Occupation of head:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage-salary worker</td>
<td>578</td>
<td>632</td>
</tr>
<tr>
<td>Farm operator</td>
<td>83</td>
<td>86</td>
</tr>
<tr>
<td>Self-employed</td>
<td>72</td>
<td>80</td>
</tr>
<tr>
<td>Unemployed</td>
<td>200</td>
<td>219</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Sex of head:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>753</td>
<td>812</td>
</tr>
<tr>
<td>Female</td>
<td>161</td>
<td>206</td>
</tr>
<tr>
<td>Recency of residence:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moved into county after Jan. 1, 1970</td>
<td>74</td>
<td>96</td>
</tr>
<tr>
<td>Lived in county since Jan. 1, 1970</td>
<td>855</td>
<td>914</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Levels of living quarters:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>816</td>
<td>859</td>
</tr>
<tr>
<td>2 or more</td>
<td>127</td>
<td>125</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>1</td>
<td>--</td>
</tr>
</tbody>
</table>
### Appendix D. Comparison between Estimated and Reported Income

#### Appendix Table 4. Comparison between 1974 estimated and reported gross income for households used in the toll demand analysis

<table>
<thead>
<tr>
<th>1974 reported income</th>
<th>1974 estimated income</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;$2,000</td>
<td>&lt;$2,000</td>
</tr>
<tr>
<td>2,000-3,999</td>
<td>3,999</td>
</tr>
<tr>
<td>4,000-5,999</td>
<td>5,999</td>
</tr>
<tr>
<td>6,000-9,999</td>
<td>9,999</td>
</tr>
<tr>
<td>10,000-14,999</td>
<td>14,999</td>
</tr>
<tr>
<td>15,000-19,999</td>
<td>19,999</td>
</tr>
<tr>
<td>30,000-49,999</td>
<td>49,999</td>
</tr>
<tr>
<td>≥50,000</td>
<td>50,000</td>
</tr>
<tr>
<td>Not ascertained</td>
<td>6,6,8,17,12,17,2</td>
</tr>
</tbody>
</table>

Table entries are the number of observations falling into each cell.
Notes

1 A main station or main telephone is a company-owned telephone instrument connected by means of an individual line or a party line directly to a central office. Each main station is given a different telephone number.

2 An extension station or extension telephone is an additional telephone instrument connected to the same circuit as the main station and subsidiary thereto.


5 Although a more logical procedure would be to first study a household's decision to rent a telephone, then its choice of grade of service, then its decision to rent an extension telephone and finally the amount it spends on long-distance calls, the demand for residential telephone services is considered in the order indicated above for the following reasons. First, toll service was the first service actually analyzed because it is believed to be more interesting than the other services. Secondly, it is easier to develop the methodology of demand analysis using toll service rather than one of the other services since the dependent variable is quantitative rather than qualitative and because much of the discussion of the demand for extensions and main stations is devoted to an explanation of the statistical techniques employed. Thirdly, variables representing tastes and preferences can be introduced into the analysis more conveniently in connection with the demand for toll service because these variables do not have as great an effect on the demand for the other services. Fourthly, the demand for extension telephones is considered before the demand for main stations since binary logit analysis can be explained with less difficulty when the dependent variable is dichotomous rather than polychotomous.

6 The studies included in this group, and a brief summary, are as follows. Using monthly data for the Wisconsin Telephone Company, Thompson and Tiao (1971) constructed a forecasting model for connects and disconnects of telephones. The dependent variable in the connects model was the sum of residential and business main telephone installa-

Soball et al. (1972) also analyzed the demand for local and toll telephone service using revenue data for all telephone companies in Canada. With these data, however, it was not possible to separate local service and toll service revenue into that obtained from residential subscribers and that derived from business users.


Waverman also applied dynamic demand models developed by Bouchhekter and Taylor (1970) to time-series data for Great Britain to estimate the demand for toll service. With these data, however, it was not possible to analyze the demand by households separately from business demand.

The following discussion is based upon the supplement provided by W. W. Betteridge in his testimony to the Federal Communications Commission, Docket No. 19129, Phase II, Washington, D. C., November 1973.


Irish does not explain how to D, where D equals unity or one, was calculated.


16 The 1960-1961 Survey of Consumer Expenditures, conducted jointly by the U.S. Department of Labor and the U.S. Department of Agriculture (1965-1966), contains data on expenditures on basic service (i.e., main stations), other telephone services (i.e., installation charges, extension telephones and vertical services), and on long-distance calls for the 14,000 individual households interviewed. In addition to these expenditure information, the survey contains data on economic and demographic characteristics of each household. Although magnetic tapes containing these disaggregated data are available, to the author's knowledge they have not been used to estimate demand functions for residential telephone services.

The unit of observation employed in conducting the Second Survey of Labor Supply in Eastern North Carolina was an economic family, defined as a group of individuals living together and pooling their income for major items of expense or an individual living alone or with others but remaining financially independent. More than one interview was obtained in housing units containing two or more economic families. Rather than combining the survey data in these cases, interviews with economic families living in multi-economic family housing units were deleted from the data set.

See Appendix A for a complete description of the long-distance information provided by CT&T.


Several dynamic demand models are available. For example, income lagged one year can be used as an independent variable along with current income. Or, current consumption can be assumed to be a geometrically declining function of current and past levels of income which, upon applying a Koyck transformation, produces an estimating equation with consumption lagged one period and current income as explanatory variables. The habit persistence or partial adjustment model developed by Nerlove (1959) produces the same estimating equation. Brown (1952) introduces consumption lagged one period directly into the model under the assumption that past levels of consumption subsume the effects of past incomes. To capture the idea that consumption of a good is influenced by a stock of habit formation, Houthaker and Taylor (1970, pp. 9-26) developed a stock-adjustment model that yields an estimating equation containing lagged consumption, lagged income, and income change as independent variables.

An estimate of total family gross income in 1972 is available for 577 of the 934 households included in the toll demand sample. However, this three-year lagged income variable was not experimented with due to the rather substantial decrease in sample size which would
...result. Long-distance records for nine months of 1974 for many of the
subscribers in the sample have also been obtained in paper form from
telephone company records. However, these 1974 data have not been
edited and placed on computer tape and, hence, cannot be analyzed at
this time. In the future it is hoped that the 1972 income data and
1974 toll information can be incorporated into a dynamic model.

28 See Appendix B for a comparison between 1972 and 1974 estimated
and reported income for those households used in the toll demand
analysis.

29 See Lee (1968, p. 480), Rain (1960a, p. 142), Puffer (1971,
p. 381), Falka (1967, p. 20), and Friedman (1957, Chapter 3).


20-36), Sichon (1976), or Pras (1956) for a detailed description
of this technique.

30 It is possible to use the data available to estimate a specific
scale if one is willing to assume, as Friedman (1952) does, that the
income scale can be set equal to the specific scale, or if one is
willing to assume that the income coefficient is equal to one for all
types of family members, as Rayner et al. (1972, pp. 60-69) do.

31 In an analysis of data for 400 residential telephone subscribers
in southern Michigan, the Survey Research Center study (1956, p. 12, see
footnote 15) discovered that a higher percentage of high toll users
than low toll users fell into the young-married, no children and older
married, no children at home life cycle classifications. However, they
were using life cycle as a proxy for relatives living away from home.
Since information on the geographical dispersion of the family
was available for the families analyzed in this study, the use of such a
proxy variable is unnecessary.

32 Although Prais and Houthakker (1951, pp. 133-134) assert that
"a valid analysis of the effects of household consumption... is
only possible if it is possible to separate out the effects of differing
family composition on consumption and will be just
as useful for purposes of predicting.

33 A study of family budget data by Houthakker (1952) found the
specific effect of children under 14 years of age to be greater for
food, clothing, and education.

34 The farm operator could be an owner-operator, an operator on
rented land, or a sharecropper. He could be working full or part time
and his dwelling unit does not have to be located on the farm property.

35 Empirical support for this hypothesis, based on analyses of other
cohorts, can be found in Burk (1950, p. 25), Lee and Phillips (1971)

36 Although Prais and Houthakker (1951, pp.
100-110), a large number of
occupational categories could be used but was not attempted in the
preliminary analysis.

37 The last category consists of family heads who had no employment
income during 1974 and includes the retired, those working on a family-
operated farm for no pay, and the involuntarily unemployed.

38 The Survey Research Center study (1956, p. 18, see footnote 15)
found the second most important use of telephone service (including
local as well as long distance) to be making and receiving business
calls or calls associated with one's work.
For example, Massei and Winnick (1960) found blacks to spend less on housing but more on food and clothing than whites.

The Survey Research Center study (1956, p. 11, see footnote 15) discovered that 90 percent of the high toll service users have some or all of their six closest friends and relatives living in another city whereas only 65 percent of the low toll users do. In addition, a variable representing the number of friends and relatives living in other places, when included in a multiple regression equation, was found to have a significant effect on the probability of a household being a high toll user.

See Appendix A for a description of the procedure used to calculate the airline distance to each relative living outside the local calling area and within the contiguous United States. The distance to relatives living outside the contiguous United States was not included in the calculations.

For a given number of a particular type of relative, as total distance to them increases toll rates increase, which may decrease the length of conversations enough to decrease toll expenditures. A priori, this is not expected to occur.

Due to Extended Area Service, calling scopes may not be just the number of telephones in the exchange in which a subscriber is located. It also includes the number of telephones in the other exchanges which comprise the subscriber's local calling area.

The calling scope variable was constructed from data contained in monthly station development reports filed from February through November 1975 by CT&T with NODC. See Appendix A for the details of this calculation.

Miscellaneous equipment, which is often referred to as vertical services, includes such things as extra bells and gongs, jacks, decorator and other fancy types of telephones, colored telephones, touch-call, etc.

Appendix A contains a detailed description of the area surveyed, the sampling technique employed, and the information gathered in these personal interviews.

See Appendix A for a detailed description of the information provided by CT&T.

Since telephone companies do not separate toll revenue into that obtained from residential customers and that collected from business subscribers, it is not possible to judge how typical the households in the sample are with respect to expenditures on toll calls. Also, the time difference between this survey and the last Census was judged to be too great to make comparisons of household characteristics meaningful.

Appendix C contains these comparisons.

An important factor affecting a respondent's decision to sign the permit slip and thus be included in the data set is believed to be the interviewer. Overall, 71 percent of the eligible families interviewed signed the permit slip. However, one interviewer obtained permission from only 15 percent of the eligible people he interviewed and three others obtained written permission from less than 50 percent of the households they surveyed. Deleting the households surveyed by these interviewers altered the empirical results very little.

There are no sampled households living in an exchange with a calling scope between 10,000 and 15,000.

Eight years or less of schooling is the omitted education category, wage-salary workers are the omitted occupation category, urban households are the omitted urbanization group, whites are the omitted race category, households with calling scopes of less than 4,000 telephones are the omitted calling scope group, and households living in the county since January 1, 1970, are the omitted recency of residence category.

For a given number of a particular type of relative, as total distance to them increases toll rates increase, which may decrease the length of conversations enough to decrease toll expenditures. A priori, this is not expected to occur.

Due to Extended Area Service, calling scopes may not be just the number of telephones in the exchange in which a subscriber is located. It also includes the number of telephones in the other exchanges which comprise the subscriber's local calling area.

The calling scope variable was constructed from data contained in monthly station development reports filed from February through November 1975 by CT&T with NODC. See Appendix A for the details of this calculation.

Miscellaneous equipment, which is often referred to as vertical services, includes such things as extra bells and gongs, jacks, decorator and other fancy types of telephones, colored telephones, touch-call, etc.

Appendix A contains a detailed description of the area surveyed, the sampling technique employed, and the information gathered in these personal interviews.

See Appendix A for a detailed description of the information provided by CT&T.

Since telephone companies do not separate toll revenue into that obtained from residential customers and that collected from business subscribers, it is not possible to judge how typical the households in the sample are with respect to expenditures on toll calls. Also, the time difference between this survey and the last Census was judged to be too great to make comparisons of household characteristics meaningful.

Appendix C contains these comparisons.

An important factor affecting a respondent's decision to sign the permit slip and thus be included in the data set is believed to be the interviewer. Overall, 71 percent of the eligible families interviewed signed the permit slip. However, one interviewer obtained permission from only 15 percent of the eligible people he interviewed and three others obtained written permission from less than 50 percent of the households they surveyed. Deleting the households surveyed by these interviewers altered the empirical results very little.

There are no sampled households living in an exchange with a calling scope between 10,000 and 15,000.

Eight years or less of schooling is the omitted education category, wage-salary workers are the omitted occupation category, urban households are the omitted urbanization group, whites are the omitted race category, households with calling scopes of less than 4,000 telephones are the omitted calling scope group, and households living in the county since January 1, 1970, are the omitted recency of residence category.

For a given number of a particular type of relative, as total distance to them increases toll rates increase, which may decrease the length of conversations enough to decrease toll expenditures. A priori, this is not expected to occur.

Due to Extended Area Service, calling scopes may not be just the number of telephones in the exchange in which a subscriber is located. It also includes the number of telephones in the other exchanges which comprise the subscriber's local calling area.

The calling scope variable was constructed from data contained in monthly station development reports filed from February through November 1975 by CT&T with NODC. See Appendix A for the details of this calculation.

Miscellaneous equipment, which is often referred to as vertical services, includes such things as extra bells and gongs, jacks, decorator and other fancy types of telephones, colored telephones, touch-call, etc.

Appendix A contains a detailed description of the area surveyed, the sampling technique employed, and the information gathered in these personal interviews.

See Appendix A for a detailed description of the information provided by CT&T.

Since telephone companies do not separate toll revenue into that obtained from residential customers and that collected from business subscribers, it is not possible to judge how typical the households in the sample are with respect to expenditures on toll calls. Also, the time difference between this survey and the last Census was judged to be too great to make comparisons of household characteristics meaningful.

Appendix C contains these comparisons.

An important factor affecting a respondent's decision to sign the permit slip and thus be included in the data set is believed to be the interviewer. Overall, 71 percent of the eligible families interviewed signed the permit slip. However, one interviewer obtained permission from only 15 percent of the eligible people he interviewed and three others obtained written permission from less than 50 percent of the households they surveyed. Deleting the households surveyed by these interviewers altered the empirical results very little.

There are no sampled households living in an exchange with a calling scope between 10,000 and 15,000.

Eight years or less of schooling is the omitted education category, wage-salary workers are the omitted occupation category, urban households are the omitted urbanization group, whites are the omitted race category, households with calling scopes of less than 4,000 telephones are the omitted calling scope group, and households living in the county since January 1, 1970, are the omitted recency of residence category.
techniques employed. An additional problem encountered in comparing cross-sectional and time-series income elasticities is that estimates of income effects obtained with household budget data are usually interpreted as long-run income elasticities while estimates obtained with time-series data are viewed as short-run income elasticities. Long-run income elasticities are larger than short-run income elasticities. A specific problem encountered in comparing Waverman's estimates with the income elasticities calculated in this dissertation is that Waverman's income elasticity shows the effects on the number of toll calls of a change in average income for the whole population, while the income elasticity estimated in this study shows the effects on toll expenditures of a change in income for households subscribing to telephone service.

66As mentioned previously, the antilogarithms of the coefficients of the household composition variables are estimates of the percentage change in toll expenditures due to an additional member in a particular age-sex group. These coefficients incorporate both income and specific effects of an additional person.

67A comparison of Model 1 and Model 3 in Table 3 reveals that exclusion of the nonsignificant variables, for which there were no strong a priori theoretical reasons for inclusion in the model, has little effect on either the estimated coefficients or the standard errors of the remaining explanatory variables. This finding, together with the low correlation coefficients shown in Table 6, suggests that the omitted variables are not insignificant due to multicollinearity and that their exclusion does not produce specification bias.

68A grade of service is particularly good proxy for local service use since the continuous use of a two- or four-party line for local messages is limited to five minutes by CT&T General Exchange Tariff, Section 28, Fourth Revised Sheet 2, filed with the NCUC on September 4, 1972.

69Even if the true standard errors of the estimated parameters in Model 4, Table 3, are 50 percent higher than the reported standard errors, which are based on the assumption of simple random sampling, all of the explanatory variables would still be significant at the 10 percent level. In addition, when Model 4, Table 3, was re-estimated with 360 dummy variables included to represent each of the segments or clusters of households, all but four of the original explanatory variables were significant at the 10 percent level. Each of these four variables is highly correlated with the location of the household and thus probably is insignificant due to multicollinearity.

70Although it is not correct to report correlation coefficients for qualitative variables, such a measure is unlikely to give an inaccurate assessment of whether there is an association between pairs of variables.

71For example, Crockett (1960, p. 310) found race to affect the income and family size elasticities of demand for food. Lee and Phillips (1971) found differences in income elasticities between urbanization groups for many items. Hamburg (1960) discovered that the Income elasticity for clothing varies with the education of the family head.

72The cost of installing an extension telephone, $1.75 if done when the main station is installed and $7.50 if installed at a later date, also cannot be investigated. Irish (1974, pp. 122-126; see footnote 12) estimated demand functions for residential extension telephones using data for individual exchanges in North Carolina. In these regressions the service connection charge variable was either insignificant or of the wrong sign. Thus, the effects of this variable are not considered to be important.

73The reader is referred to the theoretical part of Chapter 3 for a justification for including these noneconomic variables in a demand function.

74See Appendix A for a complete description of the data used in estimating the extension demand models.

75The calculation of the number of residential extension stations per residential main station for the survey area was made from information contained in the May 31, 1975, station development reports filed by CT&T with the NCUC. These reports show the number of residential main stations and extensions in each exchange at the end of each month.

76This comparison is presented in Appendix C.

77Economic problems involving discrete dependent variables are increasingly being analyzed. Examples include choice of occupation, number of transportation, number of durable goods owned, labor force participation, family size, marital status, and migration. See McFadden (1976, pp. 381-382) for a complete description of the data used in estimating the extension demand models.

78A thorough discussion of this testing procedure.

79Since the unit of observation employed in gathering the survey data was the economic family, observations on economic families living in housing units with one or more other economic families were deleted from the data set rather than attempting to combine the interview information to form one observation.

80An inside extension telephone is an extension telephone located in the same building as the main telephone. Extension telephones are provided in connection with all grades of main station telephones and service. The intention at the outset of the study was to analyze the factors affecting the exact number (i.e., 1, 2, 3 or more) of extension telephones rented. The model developed by Domencich and McFadden (1975) could be applied to study choices among more than two alternatives. However, only 25 of the households included in the extension demand sample were renting two extension telephones and only four were renting three extensions. With such few cases it would be difficult to obtain reliable statistical results concerning the factors affecting the decision to rent more than one extension station.

81The calculation of the number of residential extension stations per residential main station for the survey area was made from information contained in the May 31, 1975, station development reports filed by CT&T with the NCUC. These reports show the number of residential main stations and extensions in each exchange at the end of each month.

82This comparison is presented in Appendix C.

83Economic problems involving discrete dependent variables are increasingly being analyzed. Examples include choice of occupation, number of transportation, number of durable goods owned, labor force participation, family size, marital status, and migration. See McFadden (1976, pp. 381-382) for a complete description of the data used in estimating the extension demand models.

84As Ladd (1966, p. 170) points out, the results obtained using ordinary least squares regression to estimate a linear probability function are the same that would be produced with linear discriminant analysis, although the assumptions made concerning the distribution
the error terms and the distribution of the independent variables are different.

79 See Goldberger (1966, p. 269) for a proof of this assertion.

80 See Domenich and McFadden (1973, pp. 47-65) for a formal derivation of the binary logit probability model from a consumer's utility function.

81 If there are repeated observations for each vector of explanatory variables, equation (21) or equation (22) can be estimated by grouping observations and letting Pr(RET=1) equal the proportion of households with the same set of explanatory variables that rent an extension telephone. After performing a logit or probit transformation on this proportion, as described in Norelove and Press (1973, pp. 10-15), the equation can then be estimated using ordinary least squares regression. However, in the case where there are a large number of explanatory variables and the sample may have only one or no observations. Continuous explanatory variables are also a problem since categorizing these variables to achieve the cell frequencies required introduces, as McFadden (1973, p. 159) points out, an error-in-variable component that makes the estimator inconsistent and may make it seriously biased. For these reasons, the cell frequency approach has been rejected in favor of estimating the binary logit probability model using individual observations and maximum likelihood procedures.

82 McFadden (1973) demonstrates that, provided the data are not multilinear, the maximum likelihood estimates are unique and that the second-order condition for a maximum holds.

83 See McFadden (1973, pp. 135-139) for a proof of this last property.

84 The number of observations falling into various categories of each explanatory variable are contained in Appendix C.

85 Morrison (1972) observes that, even if the predicted probabilities are exactly equal to the true probabilities, they could not coincide with the binary (0,1) outcomes except in the extreme situation where all the true probabilities are either zero or one. For this reason he suggests that $R^2$ should be viewed as having an upper bound and provides a procedure for calculating it. Morrison recommends comparing the calculated $R^2$ with this upper bound to achieve a more meaningful evaluation of the fit achieved by a model that generates probabilities for binary outcomes. Goldberger (1973), however, views one of the assumptions underlying Morrison's analysis as leading to a special case and believes that for a binary dependent variable, as for a continuous variable, the proper upper bound on $R^2$ is unity.

86 Seriously misleading results could be produced, however, if the linear probability model is used to predict the effects of a change in an explanatory variable for households with a very small or very large original probability of having an extension telephone.

87 When the same binary logit model was estimated using reported income, a larger income coefficient was obtained. Only the results obtained using estimated income are reported here due to the better fit achieved in the extreme income ranges and because estimated income is believed to be a more accurate measure of purchasing power than reported income.

88 To see the relationship to the usual income elasticity of demand, where consumption is a continuous variable, it is only necessary to suppose that $N$ households have the choice of renting or not renting an extension telephone and the average probability that one is rented is $Pr(RET=1)$. Thus, $NPr(RET=1)$ is the number of households renting an extension initially. The effects of a unit change in income can be expressed as

\[ NPr(RET=1) * INC74EST = NPr(RET=1)Pr(RET=0), \]

and the income elasticity can be expressed as

\[ NPr(RET=1) * INC74EST = \frac{\partial Pr(RET=1)}{Pr(RET=0)} \]

which in turn is equal to equation (59). Thus, the two concepts of elasticity are virtually identical.

89 Since an income elasticity calculated using the results of estimating a binary logit probability model is not the same for each household, but depends on each household's original probability of having an extension telephone, to forecast the effects of a change in income on the number of households renting an extension telephone one should not merely extrapolate the estimated elasticity for the mean household. Instead, one should calculate an average elasticity by weighting the income elasticities corresponding to different probabilities by the relative frequency distribution of the probabilities across the population. Westin (1974, p. 10) demonstrates that elasticities based on the mean household will overestimate the true elasticity computed using a frequency distribution of probabilities. Several forecasting techniques, which take into account the initial frequency distribution of probabilities, are discussed in Chapter 6.

90 A comparison of Models 1 and 2 in Table 10 reveals that exclusion of the nonsignificant explanatory variables, for which there are no strong theoretical reasons for inclusion in the model, has little effect on either the estimated coefficients or the standard errors of the remaining variables. This finding, together with the low correlation coefficients depicted in Table 12, suggests that the omitted variables are not insignificant due to multicollinearity and that their exclusion does not produce specification bias.

91 For a detailed explanation of the likelihood ratio test see Johnson (1971, p. 460). Rau and Hill (1971, p. 256) or McFadden (1973, pp. 120-121).

92 Although the number of each type of room (i.e., dining rooms, living rooms, kitchens, bathrooms, etc.) was available from the 1973 survey, the number of bedrooms was selected as the alternative proxy variable for housing unit size for two reasons. First, these data were unedited and number of bedrooms appeared to be the most accurate of the room variables. Secondly, of all the room variables, the number of bedrooms is expected to have the highest correlation with the size of the housing unit.
At the outset of this study the intention was to analyze the factors affecting the exact number of main telephone lines rented by a family. Although two main stations per household are becoming more common, in the sample used in the empirical analysis only two households were renting more than one main telephone. Hence, this problem could not be empirically investigated.

Of the 2,113 households interviewed, all but 64 were located within the service area of CT&T. The 54 households living within the area served by the Tri-County Telephone Membership Corporation are also included in the main station demand analysis. The 10 households interviewed that were residing within the service area of the Norfolk and Carolina Telephone and Telegraph Company were excluded from this analysis.

The term "exchange" denotes a unit established by a telephone company for the administration of communication in a specified area which typically embraces a city, town or village and its environs. It consists of one or more central offices together with the subscribers' stations and lines connected thereto forming a local system furnishing communication service within that area. Within the eight-county area surveyed there were 37 exchanges, all but four of which were CT&T exchanges.

The base rate area is that section of an exchange which is within a one-mile radius from the center of the exchange. Within the base rate area the telephone company will furnish any primary grade of service at a rate that is common to all applicants and without the assessment of any charges based on distance (e.g., mileage zone charges). The main station rates applicable within this portion of an exchange are known as base rates.

Mileage zones are bands located around the base rate area within which, due to the greater distance from the central office and sparseness of development, mileage zone charges apply in addition to base rates for the classification of service furnished. Mileage zones 1A and 1B are each 0.5 mile wide; mileage zones 2A and 2B are each one mile wide; and mileage zones 3 through 10 are each two miles wide.

The term "multi-party service" is used in this dissertation to refer to two- and four-party service, although usually this term is reserved to refer to five-party or more telephone service.

Joe Domencich and McFadden (1975) for a derivation of this statistical model from utility theory, a discussion of the assumptions underlying this approach, the statistical properties of the estimator, and an application of this technique to travel demand.

If one party service is .97, between two-party and four-party rates it is .87, and between one-party and four-party prices it is .79.


The effects of urbanization on the decision to rent a main telephone cannot be studied since almost all the households living in a rural area are classified as urban households and almost the households living outside mileage zone 1B are classified as living in a rural area.

For a justification of this hypothesis, see footnote 40.

Calling scope is measured as the average size of the calling scope over the months of February, March, April, and May 1975, the period of time during which the interviewing was conducted. See Appendix A for a complete description of the calculation of this variable.

See Artle and Averoua (1973), Squire (1973), Rohlis (1974), and von Rubenau and Stahl (1974) for a theoretical discussion of the public
goods aspect of calling scope and its effect on the demand for main stations.

111 CT&T requires an advance payment or deposit from individuals of known good credit. Individuals whose credit is known to be poor are required to make an advance payment equal to the sum of the service connection charge of $15 and the local service charge for one month, plus a deposit of at least $50. Of those individuals whose credit is not known, 90 percent must pay the same advance payment and the same deposit. The remaining 10 percent of the unknown credit cases must pay the same advance payment and a deposit that depends on the anticipated amount of the credit to be extended but cannot be in excess of two-twelfths of the estimated charge for the service for the ensuing 12 months.

112 In estimating demand functions for one-party and two-party residential main stations within the base rate area of North Carolina exchanges, Irish (1974, see footnote 12) found the coefficient of the service connection charge variable to be either insignificant or of the wrong sign.

113 Perl (1975, p. 3, see footnote 14) concluded from his analysis of Census data that the effects of many explanatory variables on telephone availability are nonlinear. Also, the existence of a threshold level which varies among individuals will produce a sigmoid-shaped response curve when a group of individuals is considered.

114 The average number of hookups on two-party and on four-party lines was calculated from information contained in the May 31, 1975, station development reports filed by CT&T with the NCUC. The average number of hookups on two-party lines is equal to the total number of two-party main stations (business plus residential) in the exchange divided by the number of two-party lines in the exchange. The average number of hookups on four-party lines equals the number of four-party main stations in the exchange divided by the number of four-party lines. See Appendix A for the details of these calculations.

115 Since the appropriate unit of analysis is the household, surveys for families living in multi-family housing units were deleted from the data set rather than attempting to combine the data to form one observation. See Appendix A for a complete description of the data used in these analyses.

116 Since March 5, 1973, two-party residential main station telephone service has not been available to new subscribers or to existing subscribers who move into a base rate area. A residential subscriber who was renting a two-party main station prior to March 5, 1973, might have had a two-party line on the day interviewed as long as there were still two stations connected to his line. When only one station is connected to a two-party line within the base rate area, the telephone company reclassifies the service as one-party and applies the one-party rate.

117 The available evidence indicates that this is a reasonable assumption. Using data supplied by CT&T in response to the NCUC Supplemental Order, Group A, Item 2b (NCUC Docket No. P-7, Sub 660, Raleigh, June 3, 1975) and the company's December 31, 1974, station development reports filled with NCUC, an estimate of the number of residential subscribers in the base rate area of each exchange in the eight-county survey area was calculated. A comparison of these figures with company-provided data on the number of residential two-party lines in the base rate area of these exchanges on December 31, 1974, reveals that only 0.75 percent of the residential subscribers in this section of an exchange were renting two-party lines. This figure would be expected to be even lower at the time of the survey due to continuing efforts to upgrade service. In addition, of the families surveyed that were living in a base rate area and for whom grade of service information has been obtained from the telephone company, only 3.44 percent were renting a two-party main station on June 30, 1975.

118 If a rural or urban family reported that there was no telephone installed in its home, the validity of this response was checked by consulting telephone books for the survey area and, in some cases, by calling directory assistance.

119 For statistics for various years on the percent of households in the United States with a telephone, see U. S. Bureau of the Census (1975, p. 515).

120 The difficulties involved in using an income elasticity evaluated for the mean household for forecasting have been pointed out in Chapter 4. Alternative forecasting techniques are discussed in Chapter 6.

121 The difficulties encountered in comparing estimates obtained in different studies are numerous and are discussed in footnote 62.

122 Throughout this summary when the toll or extension demand model is being referred to, a variable is termed significant if its estimated coefficient is significantly different from zero at the 5 percent level.

123 Stern (1965, p. 27) cites a case where, during a one-month period after the Chesapeake and Potomac Telephone Company of West Virginia was authorized to put into effect an increase in main station rates in 1955, 250 services (business as well as residence) were downgraded at customer request stating the higher rates as the reason. After the next increase, in 1957, 104 services were downgraded.

124 There is no charge made by CT&T for changes in grade of service.


126 See Lansing and Morgan (1971, p. 139).

The date to which all the local service information corresponds is June 20, 1975, the day on which this information was extracted from the computerized billing records maintained at CT&T company headquarters in Tarboro, North Carolina.

The name and address to which the telephone bill is sent was compared with the same and address for the household interviewed to insure that the telephone information obtained corresponded to the household interviewed. If it did not, an investigation was conducted to determine the source of the discrepancy and appropriate corrections were made, if necessary.

Exchange Area Service Maps are filed with the NCOG and made a part of CT&T's Local Exchange Tariff, NCOG Docket No. 7-7, Sub 582, Raleigh, March 5, 1973.

Price and calling scope information for households located within the service area of the Tri-County Telephone Membership Corporation, a nonregulated telephone company, were obtained in a September 16, 1973, telephone conversation with Cecil Smith, manager.

Table entries are the number of observations in each data set falling into each category of each variable.