2
MASTER

Ohio

Time-of-Use Electricity Price Effects :

Prepared for the U.S. Department of Energy
Office of Utility Systems

S.B. White • C.A. Clayton
B.V. Alexander • D.P. Duncan

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

MASTER

TIME-OF-USE ELECTRICITY PRICE EFFECTS OHIO

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Prepared for
The Department of Energy
Office of Utility Systems

By
S. B. White
C. A. Clayton
B. V. Alexander
D. P. Duncan

December 1979

Contract No. CR-04-80220-00/RTI Project No. 41U-1542/00-12F.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

PAGES ii to iv
WERE INTENTIONALLY
LEFT BLANK

ACKNOWLEDGMENTS

This report was prepared by the Research Triangle Institute (RTI), Research Triangle Park, North Carolina, pursuant to Contract No. CR-04-80220-00 with the U.S. Department of Energy (DOE). The statements, findings, conclusions, and recommendations presented in this report do not necessarily reflect the views of the Department of Energy. This report has benefited from the many helpful comments and suggestions of DOE staff members Steven Mintz (Project Officer) and Tom Howze.

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK

ABSTRACT

The Ohio Rate Demonstration Project was conducted to observe the effect of a time-of-use (TOU) electrical rate on household electricity consumption. Under time-of-use rates higher prices are charged for electricity used during those parts of the day when demand and system costs are higher, and lower prices are charged during those parts of the day when demand and system costs are lower. In the Ohio project 160 customers of the Dayton Power and Light Company (100 on a seasonal time-of-use electricity rate and 60 on the control rate) became part of an 18-month study from June 1976 to November 1977. The experimental rate schedule used two rating periods during weekdays and a single rating period on weekends and holidays. All participation in the study was voluntary. Participants did not receive any direct compensation.

A comparative analysis, by calendar months, of the electricity consumption patterns of customers on the TOU rate and a control group on the existing (declining block) rate showed that customers on the TOU rate used less electricity than the control group during the six-hour period of peak prices and used correspondingly more electricity during the base period of low prices. As a result of these offsetting effects, overall consumption was unaffected. Tests of hypotheses showed the peak-period decreases (ranging from 21 to 38 percent on a monthly basis) to be statistically significant. On the other hand, the base-period increases were not significant (at the 5 percent level) because of much larger customer-to-customer variation in base period usage.

Other significant findings were that the diversified demand of the residential class at the time of system peak was, in the two months examined (August and September 1976), significantly lower for the experimental group. The residential class load factor based on that diversified demand was significantly higher for the experimental group during August. At the time of system peak, diversified demand was lower by 29 to 40 percent and the associated load factor was 135 to 161 percent of that under the control rate.

These results apply to the group of residential customers who met all of the following conditions: a) had a 12-month billing history (as of April 1975), b) did not live in an apartment house or trailer, c) were not employed by Dayton Power and Light Company, d) were not on a frozen rate, e) had annual usage greater than 6,000 kWh for period April 1974 -April 1975, f) volunteered to participate in this study, g) did not plan to move during the course of the study, and h) could be metered (conveniently). It is estimated that 42 percent of the 337,000 residential customers failed to meet one or more of the first five conditions. Of the remaining customers about 80 percent failed to meet one or more of the remaining conditions. This leaves approximately 39,000 as the population to which inferences can be drawn from the sample results.

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page</u>
1	INTRODUCTION	1
	1.1 Objectives and Applications of Findings	1
	1.2 Summary Description of the Experimental Project . .	2
	1.3 Summary of Findings	3
	1.4 Limitations of the Demonstration and Qualification of Analytical Results	6
2	PROJECT CHARACTERISTICS	9
	2.1 History and Major Features of the Project	9
	2.2 Description of Rate Schedules	10
	2.3 Customer Participation Payments and Interviews . .	10
	2.4 Design of the Sample	13
	2.5 Rejections during Sample Selection	14
3	COMPARATIVE ANALYSIS	17
	3.1 Analytical Approach	18
	3.2 Selection of Parameters for Comparison	19
	3.3 Preparation of Data for Comparative Analysis . . .	23
	3.4 Comparative Analysis Results	28
	3.5 Average Daily Load Curves	32
REFERENCES		41
APPENDIXES		
A	DESCRIPTION OF DATA PROCESSING AND ANALYSIS METHOD . . .	A-1
B	STATISTICAL ESTIMATION AND TESTING PROCEDURES USED IN COMPARATIVE ANALYSIS	B-1
C	LISTING OF CUSTOMERS DELETED IN MANUAL EDIT	C-1
D	COMPARISON OF EXPERIMENTAL RESPONSES UNDER TOU RATE AND CONTROL RATE	D-1

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK

LIST OF TABLES

<u>TABLE</u>		<u>PAGE</u>
1-1	Summary Features of the Ohio Rate Demonstration Project .	3
2-1	Specification of Rate Schedules Used in the Ohio Rate Demonstration Project	11
2-2	Criteria for Inclusion in the Demonstration	16
3-1	List of Parameters Estimated in Comparative Analysis . . .	20
3-2	Summary of Data Reduction	24
3-3	Number of Customers Flagged by Month by Test Group in the Automated Edit	26
3-4	Number of Customers Deleted by Month by Test Group in the Automated and Manual Edits	27
3-5	Customers Providing Usage Data for Analysis	27
3-6	Differences between Control and Experimental Groups for Values Deemed Significant	29
3-7	Average Daily Load Curves	33
A-1	Form of Records in I-1	A-7
A-2	Form of Records in File U-2	A-8
A-3	Form of the U-5 Data File	A-11
A-4	List of Parameters Estimated in the Comparative Analysis .	A-15
D-1	List of Parameters Estimated in the Comparative Analysis .	D-2
D-2	Comparison of Experimental and Control Rate Groups, August 1976	D-3
D-3	Comparison of Experimental and Control Rate Groups, September 1976	D-4
D-4	Comparison of Experimental and Control Rate Groups, December 1976	D-5
D-5	Comparison of Experimental and Control Rate Groups, January 1977	D-6

LIST OF TABLES (Continued)

<u>TABLE</u>		<u>PAGE</u>
D-6	Comparison of Experimental and Control Rate Groups, July 1977	D-7
D-7	Comparison of Experimental and Control Rate Groups, August 1977	D-8
D-8	Comparison of Experimental and Control Rate Groups, September 1977	D-9

LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1-1	Average Weekday Load Curves, August 1976	5
1-2	Day of System Peak Load Curves, August 1976.	5
2-1	Schematic of Ohio Experimental Rate Schedule	12
3-1	Average Daily Load Curve for Control and Experimental Rate Schedule, August 1976	34
3-2	Average Daily Load Curve for Control and Experimental Rate Schedule, September 1976	35
3-3	Average Daily Load Curve for Control and Experimental Rate Schedule, December 1976	36
3-4	Average Daily Load Curve for Control and Experimental Rate Schedule, January 1977	37
3-5	Average Daily Load Curve for Control and Experimental Rate Schedule, July 1977	38
3-6	Average Daily Load Curve for Control and Experimental Rate Schedule, August 1977	39
3-7	Average Daily Load Curve for Control and Experimental Rate Schedule, September 1977	40
A-1	Overview of Analysis System	A-3
B-1	Population and Sampling Structures	B-4

CHAPTER 1: INTRODUCTION

In 1975 the Federal Energy Administration (now part of the U.S. Department of Energy) initiated experiments in eight states to determine how household consumption of electricity is affected by time-of-use pricing.* Funding for the projects was provided by the Federal Energy Administration, participating utilities, and state governments. In the fall of 1977 the Department of Energy (DOE) contracted with the Research Triangle Institute (RTI) to perform a comprehensive, standardized analysis on the data from the residential components of these studies. This report on the Ohio Rate Demonstration Project is one of a series of final reports which will be published under this contract. (For a listing of these reports and a summary of the standardized analytical approach, see the RTI Analytical Master Plan [RTI 1978].)

1.1 Objectives and Applications of Findings

The objective of the analyses reported here is to determine how electricity consumption by residential customers on a time-of-use (TOU) rate differs from electricity consumption by residential customers under the existing rate. The findings of this study provide information which can be used by both regulators and utility system planners to: (1) project changes in total capacity and energy requirements if the TOU rate schedule were to be implemented, (2) project possible changes in the optimal mix of generating capacity, and (3) determine whether the potential benefits of altered load patterns under the TOU rate offset the potential additional metering costs of implementing such a rate.

*These eight projects were in Arizona, Arkansas, Connecticut, California (Los Angeles), New Jersey, Ohio, Wisconsin, and Vermont. Subsequently, six projects were added in California, North Carolina, Oklahoma (Edmond), Puerto Rico, Rhode Island, and Washington.

1.2 Summary Description of the Experimental Project

Initiated in June 1976, the Ohio Rate Demonstration Project was an 18-month controlled experiment in which 100 Dayton Power and Light Company (DP&L) residential customers were placed on a seasonal time-of-use (TOU) rate schedule. The rate utilized two rating periods during weekdays and a single rating period during weekends and holidays. Sixty customers on the existing (declining block) rate served as a control group. Table 1-1 summarizes characteristics of the Ohio Project.

The population of customers from which the sample was drawn consisted of all residential customers during the period from April 1974 to April 1975 with the exception of the following: (1) customers without a 12-month billing history, (2) customers living in bulk-metered residences such as apartment houses and trailers, (3) customers employed by the utility company, (4) customers on frozen rates, and (5) customers with less than 6000 kWh annual usage. These exclusions represent 42 percent of the total residential customers in the DP&L service areas--a rather large number of exclusions prior to sampling.

This sampled population (approximately 196,000) was stratified into 100 groups according to annual consumption during the period from April 1974 to April 1975. Each stratum had an equal number of customers. Simple random sampling was employed within each stratum to obtain one customer to participate in the experimental group. A similar procedure was followed for the control group, except that only 60 strata were formed.

All participation in the study was voluntary. Participants did not receive any direct compensation.

This report covers analyses made on the following selected months: August, September, and December 1976 and January, July-September 1977.

Table 1-1. Summary Features of the Ohio Rate Demonstration Project

Participating Agencies/Utilities: U.S. Department of Energy, Ohio Public Utilities Commission, Dayton Power and Light Company (DP&L).

Customer Participation: voluntary.

Study Period: test period, 18 months, June 1976 to November 1977.

Metering: thirty-minute kWh meters.

Rates: one seasonal time-of-use (TOU) rate with two daily rating periods on weekdays--peak (P) and base (B). A single rating period existed on weekends and holidays--(B).

Number of Customers: 100 customers on the seasonal TOU rate, 60 in control group.

Incentives/Revenue Neutralizers: revenue neutral TOU rate.

These months were chosen from the test period data because they cover the principal period of interest for the winter and summer seasons.

For each month, 20 population parameters were estimated for the customers on the TOU and control rate schedules. Parameters included such quantities as usage during peak and base rating periods, noncoincident maximum demand, and diversified demand at time of system peak. Table 3-1 in chapter 3 provides a complete list of the parameters that were estimated. Comparisons of these estimated population variables were made between the control and experimental groups to see if the differences between the consumption patterns of the two groups were statistically significant. A series of average daily load curves utilizing various averaging periods (all weekdays and all weekend days, for example) were also developed for each month.

1.3 Summary of Findings

The findings summarized on the following pages are based on the comparative analysis method developed by the Research Triangle Institute. A

discussion of the analytical methodology is given in Volume I of the Analytical Master Plan (RTI 1978) and in Appendixes A and B of this report. In this analysis the electricity consumption patterns of the sample were characterized by several parameters. Then differences in estimates of these parameters between the experimental and the control group were tested for statistical significance. In addition to the comparative analyses, load curves were constructed to provide visual comparisons of the differences in the electricity consumption patterns of the two groups. The graphs in figures 1-1 and 1-2 offer two examples of representative load curves.

The results of the study are summarized below. They are considered applicable only to the population of residential customers actually represented by the sample.* The reasons for the limitations on application of these analytical results to a larger population are explained in the next section.

Peak Period. During the six-hour period of high prices, customers on the TOU rate used significantly less electricity than customers on the existing (declining block) rate. The reduction in usage was relatively consistent over the time period examined (which encompassed all seasons). Monthly reduction in usage during the peak period ranged from 21 percent to 38 percent or from 1.52 to 2.62 kWh per customer per day.

Base Period. Some shifting of usage from the peak to the base period occurred, as evidenced by an increase in usage during the base period in seven of the eight months examined. Shifting was more pronounced during the winter months. Statistically, these increases were not significant. However, such small shifts in usage are not detectable with the present sample sizes and the

*This point cannot be overemphasized. The population of inference is estimated to contain approximately 39,000 customers. This represents 20 percent of the sampled population or 12 percent of residential customers in the DPE service area.

Figure 1-1. Average Weekday Load Curves, August 1976

Figure 1-2. Day of System Peak Load Curves, August 1976

large variation in usage from customer-to-customer observed during the base period.

Total Consumption. Total monthly consumption for the experimental group did not differ significantly from that of the control group. This is attributable to the fact that the decrease during the peak period was offset by an increase during the base period although the latter was not statistically significant.

Diversified Demand. For the two months in which system load data were available (August and September 1976), the diversified demand at the time of system peak for the experimental group was significantly less than that for the control group. The associated class load factor for the experimental group was significantly higher in the month of August.

Usage during the Hour preceding/following the Peak Period. During the one-hour periods preceding and following the peak period, customers on the TOU rate schedule did not use significantly more electricity than customers on the existing rate schedule.

Noncoincident Maximum Demand. There is little evidence to indicate that the experimental and control groups differ with regard to noncoincident maximum demand.

1.4 Limitations of the Demonstration and Qualification of Analytical Results

At the time the sample was drawn, some 337,000 residential customers lived in the Dayton Power and Light Company Service area. Approximately 42 percent or 141,000 customers were excluded before the sample was drawn for the reasons stated in section 1.2. During the sampling process, a very large number (about 80 percent) of the customers selected to participate in the study were not included for various reasons, e.g. meter installation problems, refusal to participate, and moving plans. Based on the combined number of

rejections for the experimental and control groups, customers in the final sample represent about 39,000 residential customers. This group is defined as the effective population and constitutes approximately 20 percent of the sampled population and 12 percent of all residential customers.

In our opinion, the effective population is the only population to which valid inferences can be drawn from the sample results. Inferences to the sampled population or to any population larger than the effective population require the assumption that excluded customers will respond to the experimental rate in the same manner as those represented in the sample. Since this assumption was not made, all test results shown herein are applicable only to the effective population.

Even for the effective population, a number of additional qualifications apply to the results. They may be summarized as follows:

- The results strictly apply only to the represented portion of the DP&L residential class as of April 1975, the time at which the sample was drawn. Basically this group consists of customers using more than 6,000 kWh annually, who would volunteer for a TOU rate and could be metered.
- The results strictly apply only to comparisons of the particular TOU and control rate schedules incorporated in the experiment.
- A small number of customers were removed from the original sample by edit procedures prior to analysis because of data anomalies (see Appendix C). The extent to which this affected the population of inference is unknown.
- During the experiment some customers were dropped because they moved. To extrapolate observed load patterns from the remaining sample to the initial effective population requires the assumption that such attrition does not introduce systematic bias, e.g., by representing an increasingly sedentary population in the later months of the test period.
- Participants knew that the experiment was limited to a single 18-month period after which they would return to the existing rate schedule. This may have discouraged them from investing in appliances and home improvements that are financially attractive under TOU rates.

- . Finally the results reported in this study are primarily short-term responses of a small portion of the Dayton Power and Light Company customers living in the service area at the time the sample was drawn. The information provided by this experiment does not enable statistical inferences to be made about the long-term responses of some future customers to any TOU rate schedule. Furthermore, it does not enable such inferences about either short-or long-term responses to a TOU rate different from that used in this study; therefore, "real-world" application of these results require assumptions or "best guesses" that are not statistically testable.

CHAPTER 2: PROJECT CHARACTERISTICS

This chapter describes design characteristics of the Ohio Rate Demonstration Project. The first section summarizes major features of the experiment, including such items as the study period and the number of customers involved. The second section describes the rate schedules used in the project and section three states the terms of customer participation. Section four summarizes the design of the sample and the final section reviews the rejections during sample selection.

2.1 History and Major Features of the Project

Under the sponsorship of the Department of Energy, the Ohio TOU pricing experiment was conducted through the joint efforts of the Public Utilities Commission of Ohio and the Dayton Power and Light Company.

The experiment began in June 1976. It was an 18-month controlled experiment in which 100 Dayton Power and Light Company (DP&L) residential customers were placed on a seasonal TOU rate schedule. The rate utilized two rating periods during weekdays and one rating period during weekends and holidays. Sixty customers on the existing rate served as a control group. Two personal interviews were conducted with customers in the experiment. The purpose of the first contact was to ask customers to participate in the experiment. A follow-up interview was held later in the project to gather demographic data.

From January 10, 1977 through April 10, 1977 DP&L employees were on strike. During that period data tapes were picked up regularly for all households in the experiment. However, it was not possible for the Ohio Project staff to edit the data tapes or update the history file.

The TOU experimental billing was temporarily suspended due to DP&L's inability to process the data. For a short time, customers were billed on the

existing rates, but were informed that their bills would be adjusted to TOU rates as soon as possible.

2.2 Description of Rate Schedules

Two rates--control and experimental--were included in the Ohio Project. The control was the existing rate (declining block) at the time the study began. The experimental rate schedule was a TOU rate with two daily rating periods during weekdays and one daily rating period on weekends and holidays. The rating period also varied by season. During summer and fall weekdays, the TOU rate schedule had a six-hour peak period from 12:00 noon until 6:00 P.M. The winter and spring weekday experimental rate also had a six-hour peak from 1:00 P.M. until 7:00 P.M. All other hours were off-peak.

There were two seasonal variations of the TOU rate. The summer and winter months had the same peak and base prices. The fall and spring peak and base prices were also the same.

Detailed specifications of the control and experimental rate schedules are given in table 2-1. As noted in table 2-1, the control rate changed twice during the study period. Figure 2-1 shows the daily rating periods of the experimental rate schedule as it was applied on weekdays during the various seasons of the experimental period.

2.3 Customer Participation Payments and Interviews

Customer participation was voluntary for members of both the experimental and control groups. No direct compensation was given to the study participants. Further details about how the sample was drawn and assignments made to the TOU and control rate schedules are given in the next section.

Table 2-1. Specification of Rate Schedules Used in the Ohio
Rate Demonstration Project

<u>Experimental Rate</u> Time	<u>Hours</u>		<u>Price (¢ per kWh)</u>	
	P	B	P	B
Summer Weekday	12-18	18-12	9.1	0.4
Fall Weekday	12-18	18-12	3.1	0.4
Winter Weekday	13-19	19-13	9.1	0.4
Spring Weekday	13-19	19-13	3.1	0.4
All Weekends		00-24		0.4

Summer: June 15 - September 15

Fall: September 16 - November 14

Winter: November 15 - February 15

Spring: February 16 - June 14

\$6.00 per month customer charge

P = peak; B = base

The following six holidays are treated as weekend days: New Year's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving, and Christmas Day.

Control Rate

First 30 kWh or less*	- \$2.80
Next 170 kWh	- 3.59 cents per kWh
Next 550 kWh	- 2.30 cents per kWh
Additional kWh	- 1.30 cents per kWh
First 30 kWh or less†	- \$2.81
Next 170 kWh	- 3.623 cents per kWh
Next 550 kWh	- 2.333 cents per kWh
Additional kWh	- 1.333 cents per kWh

*Effective 10/23/76

†Effective 2/18/77

*On all weekends and holidays the rate for all hours of the day was 0.4¢ per kWh.

Figure 2-1. Schematic of Ohio Experimental Rate Schedule

Customers participating in the experiment were interviewed personally on two occasions. Prior to the start of the study, an interview was held to determine if homes were suitable for meter installation and to request permission for the installation of the monitoring equipment. At that time housing, appliance stock, income, and demographic data were obtained by questionnaire. During the initial interview, the TOU experimental customers also received information on the cost savings associated with conservation and the types of activities which might be shifted to off-peak hours. In addition, since the initial bill was rather complicated, it was hand delivered by the service representative and the various components explained to the customers. These customers also received a letter advising them when a change in peak hours was about to occur.

Interviewers for this project were company service representatives briefed by their project supervisors prior to their contact with customers. They had no other special training. Toward the end of the study a combined follow-up and updated demographic questionnaire was administered to customers in the experiment.

2.4 Design of the Sample

Sample customers were selected from the group of residential customers in the DP&L service area that remained after the following exclusions were made: (1) customers without a 12-month billing history for April 1974 to April 1975; (2) bulk-metered residences; (3) company employees; (4) residents on "frozen" rates; and (5) customers with less than 6000 kWh annual usage, which accounted for about 40 percent of the customers. The group that was sampled numbered approximately 196,000 customers.

The initial sample of 100 customers for the experimental group was obtained by stratifying all eligible customers as described above according to

annual consumption during the period from April 1974 to April 1975. Each stratum contained the same number of customers. In order to obtain one participant per stratum for the final sample, six customers were actually selected within each stratum, using systematic sampling with a random start. For the control group, 100 strata were formed with an equal number of customers in each stratum. However, these strata covered all customers regardless of annual usage--that is, customers with a prior usage of less than 6,000 kWh were not excluded. Under this stratification scheme only 60 of the 100 strata involved customers with annual usage greater than 6,000 kWh. In order to have a valid control group sample (i.e., a sample compatible with the experimental group sample), only data from the 60 customers in the 60 high-usage strata were utilized. The initial selection of six customers per stratum was made simply to avoid having to redraw additional samples.

2.5 Rejections during Sample Selection

In the process of enrolling customers selected in the sample for inclusion in the study, a large number were rejected for one of several reasons. For example, in the case of the experimental group a total of 551 customers were contacted in order to obtain the desired sample of 100. For the control group, a total of 413 customers were contacted to obtain the required sample of 100. Some of the major reasons for the high rejection rate include: refusal to participate, moving plans, and meter inaccessibility. A more detailed breakdown of the number of customers by reason of rejection was not made available to RTI.

Assuming a uniform rejection rate for the control group across all 100 strata, it is estimated that 248 (i.e. 60 percent of 413) customers were screened in order to obtain the desired sample size of 60 customers with prior

annual usage greater than 6000 kWh. (This assumption was necessitated by the lack of data on rejection on a stratum-by-stratum basis). Over both the experimental and control groups, it is estimated that a total of 799 (= 551 + 248) customers were screened to furnish the total sample of 160 participants (100 on the experimental rate and 60 on the control rate). The combined rejection rate for the two groups is therefore estimated to be about 80 percent. Applying this rejection rate to the sampled population of 196,000 gives an estimated 39,000 customers in the population actually represented in the sample. This represents the population for which valid inferences can be made from the sample results. The estimated size of the universe, sampled population and effective population for the Ohio study are given in table 2-2.

Table 2-2. Criteria for Inclusion in the Demonstration

Inclusion Criterion	Percent Included	Number of Customers
(1) Universe: Estimated number of residential customers in DP&L service area - April 1975	100%	337,000
(2) Sampled population: Estimated number of residential customers in sampled population (excludes customers without a 12-month billing history, living in bulk-metered residences, working for DP&L, on "frozen" rates, or having less than 6000 kWh annual usage).	58% of (1)	196,000
(3) Effective population: Estimated number of residential customers in population represented by sample (includes those customers not excluded during sample selection due to metering problems, moving plans, and refusal to participate).	20% of (2)	39,000

CHAPTER 3: COMPARATIVE ANALYSIS

Since only one experimental rate schedule was used in the Ohio study, this chapter compares electricity consumption of customers on an experimental TOU rate with customers on a control rate. The results can be statistically projected only to those DP&L customers who volunteered to participate in the experiment and would not have been excluded for the specific reasons described in section 2.4 and in section 2.5. The results of the study are summarized in the following paragraphs.

Peak Period. During the six-hour peak period, customers on the TOU rate used significantly less electricity than the control group customers on the existing (declining block) rate. The reduction in usage was relatively consistent over the seven months examined (August, September, and December 1976, and January, July-September 1977). Monthly reduction in usage during the peak period ranged from 21 to 38 percent or 1.73 to 2.62 kWh per customer day.

Base Period. Some evidence of a shift in usage from the peak to the off-peak period exists particularly during the winter months. In six of the seven months examined, the estimated mean usage in the base period under the experimental rate was higher than for customers under the existing rate. The monthly increases were not statistically significant because of large customer-to-customer variation in base-period usage.

Total Consumption. Total monthly consumption for the experimental group did not differ significantly from that of the control group. This is because the decrease during the peak period was offset by an increase during the base period even though the latter was not statistically significant.

Diversified Demand. For the two months in which system load data were available (August and September 1976), the diversified demand at the time of

system peak for the experimental group was significantly less than that for the control group. The class load factor based on demand at the time of system peak (the ratio of average hourly consumption during the month to hourly consumption at time of system peak) for the experimental group was significantly higher than the control group in August (0.997 as compared to 0.619).

Usage during the Hour preceding/following the Peak Period. During the one-hour periods preceding and following the peak period, customers on the TOU rate did not use significantly more electricity than customers on the existing rate schedule.

Noncoincident Maximum Demand. There is little evidence to indicate that the experimental and control groups differ with regard to noncoincident maximum demand.

The sections which follow describe the analysis and these overall findings in detail. The first two sections describe the analytical approach for the comparative analysis and identify the population parameters to be estimated. This is followed by a description of the preparation of the data for the comparative analysis. Then in section 3.4, the results of tests of hypotheses on these parameters are reported. The final section describes the four types of load curves constructed for the study and provides plots for selected months.

3.1 Analytical Approach

The comparative analysis involves a statistical comparison of responses observed under the experimental and control rate schedules. The objective of the analysis is to determine if the TOU rate causes significant changes in usage patterns. The analysis used to determine whether a response under one rate schedule differs significantly from the response under another rate form

rests on the assumption that a definable population exists about which inferences from samples can be drawn. Weights are used to expand the sample to a finite population.

For this study, stratification was ignored and the data were analyzed as a simple random sample. This approach was taken for several reasons. Single observations within strata would require extensive collapsing of strata to estimate variances. High rejection rates led to a relatively small population to which results are applicable. In addition, the Department of Energy data file did not contain stratification codes and the formation of a new set of strata would require additional programming efforts. Finally, information on sample rejections by stratum was not available. Thus, estimated means shown herein are actually sample means with application only to the effective population. The standard errors take into account the size of the effective population and reflect the fact that the number of customers screened to obtain the desired sample size is a random variable. (See Equation (B-2) of Appendix B with H, number of strata, equal to one.)

3.2 Selection of Parameters for Comparison

Five summer months August and September 1976 and July -September 1977 and two winter months December 1976 and January 1977 were selected for analysis. For each of these seven months, the applicable parameters of table 3-1 were estimated for the effective* population for the control and experimental rates (that is, assuming these rates were applied to the effective population). The twenty parameters shown in table 3-1 were selected because they are important

*As noted earlier, based on combined rejections from the experimental and control groups, customers in the final sample represent about 39,000 residential customers. This group is defined as the effective population and constitute approximately 20 percent of the sampld population.

Table 3-1. List of Parameters Estimated in Comparative Analysis
(Estimated per customer)

	Parameter Number	Description
Average for Day of Monthly System Peak	1	kWh - peak
	2	kWh - intermediate
	3	kWh - base
	4	kWh - during 1-hr. preceding peak period
	5	kWh - during 1-hr. following peak period
	6	Total kWh
Average for All Days of Month	7	kWh - peak
	8	kWh - intermediate
	9	kWh - base
	10	kWh - during 1-hr. preceding peak period
	11	kWh - during 1-hr. following peak period
	12	Total kWh
Average per Customer	13	Diversified demand at time of observed monthly system peak
	14	Noncoincident maximum demand
Average for Subset of Hours during Month	15	Average hourly kWh consumed during all hrs. of observed system load greater than 80% of <u>annual</u> system peak
	16	Average hourly kWh consumed during all hrs. of observed system load greater than 90% of <u>annual</u> system peak
	17	Average hourly kWh consumed during all hrs. of observed system load greater than 95% of <u>annual</u> system peak
Ratio	18	Proportion of total kWh used during peak period on peak day [$\#1/\#6$]
	19	Proportion of total kWh used during peak period on average day [$\#7/\#12$]
	20	Class load factor based on diversified demand at time of observed monthly system peak ($\#12/\#13 \div 24$)

and meaningful response measures for comparing the control and experimental rates.* For a more technical listing of these parameters, see table A-4. The objective was to determine, for the population of experimental customers, whether the level of response for a particular parameter was a result of the rate structure being used.

Parameters Defined. Several general comments may aid the reader in understanding these parameters. Parameters 1-6 of table 3-1 measure average usage per customer during certain hours of only one day of the month, the day of system peak. For example, parameter 1 measures usage during the hours defined as the peak period on that day and parameter 6 measures total usage over all hours. Parameter 2 is not applicable to the Ohio rate schedule since it involves only two TOU pricing periods--peak and base. Parameters 4 and 5 each measure usage during only a single hour on the peak day; the hours before and after the peak period, respectively. The purpose of these variables is to determine whether the TOU rates cause a surge in demand at those times. Parameters 7 and 9-12 are identical to parameters 1 and 3-6 except that they are averaged over all days of the month.

The word "demand" is usually defined in electricity load research as the instantaneous load (in kilowatts) that a customer or customer group imposes on the utility system. "Diversified demand" is the average of two or more customers' "demands" at any instant. The word "diversity" connotes the extent of differences among individual customers' load curves. For example, if all customers had identical load curves, there would be no diversity and the average "maximum diversified demand" for the group would be the sum of their

*These parameters represent the end product of discussions held with regulatory and utility representatives during a two-day DOE-sponsored conference in Washington, D.C. in December 1977.

individual maximum demands, also called "noncoincident maximum demand," divided by the number of customers in the group. Given these definitions, it follows that parameter 13 is the average kWh usage per customer during the single hour in each month when the largest hourly demand was observed on the entire utility system. By contrast, parameter 14, noncoincident maximum demand, is the estimated average over all customers in the effective population of their individual maximum hourly kWh usages. This was computed as the weighted average of the maximum that was observed for each sample customer within each month.

Parameters 15-17 provide usage characteristics on a per customer basis averaged over subsets of hours identified by the magnitude of the observed load on the utility system. For example, parameter 16 averages hourly kWh consumption over those hours in the month when system load exceeds 90 percent of the annual system peak.

Parameters 18 and 19 are the proportion of total usage during the peak period on the day of system peak and for all days in the month, respectively. For example, during August 1976 the customers on the TOU rate used 20.6 percent of their total usage during the peak period on the peak day and 14.4 percent on the average day. These contrasted with corresponding estimates of 32.7 percent and 22.3 percent for the control group.

Load factor, the ratio of the average load to the peak load, can be measured by system or class. Class load factor, parameter 20, is the ratio for the residential customers of average hourly usage to average usage during the hour of system peak within each month. The numerator in this ratio is parameter 12 divided by 24 and the denominator is parameter 13. For example, during August 1976, the average hourly usage of the control group was 1.288 kWh and the average usage during the hour of system peak was 2.080 so the class load factor based on the latter was 0.619 ($=1.288/2.080$).

The parameters in table 3-1 were estimated for customers on the control and experimental rates for seven months covering the eighteen-month test period. Purposely excluded were those months in which rating periods changed on a billing month basis and those months during which the Dayton Power and Light Company employees were on strike. Note that parameters 1-6, 13, 15-18 and 20 are estimated only for August and September--months for which system load data were available. A test of the hypothesis that the true population parameters for the control and experimental rate schedules do not differ was then conducted for each of the twenty parameters. A rejection of the hypothesis suggests evidence of a "real" difference between the two population parameters and that this difference is attributable to the rate structures.

3.3 Preparation of Data for Comparative Analysis

Prior to beginning the comparative analysis, an examination of system load data was made to identify time frames for the comparative analysis and to formulate explicit parameter definitions. Observations of system load data under the control rate schedule provided the following information for selected months: hourly system load, average hourly system load, maximum hourly system load, system load factor, and day and hour of maximum hourly system load. Explicit definitions of the time frames for many of the parameters described in Section 3.2 depend on such system load characteristics as the value of annual system peak (1793 MW for calendar year 1976) and the value and time of occurrence of the monthly system peak. For example, parameter 15 requires an aggregation over those hours in the month in which system load exceeds 80 percent of the annual system peak (the number of such hours was 92 in August and 16 in September). The results of the system load data reduction are given in table 3-2. Of the calendar months selected for

analysis, system load data were available in the DOE data bank only for two months--August and September 1976.

Table 3-2. Summary of Data Reduction

	1976			1977			
	Aug	Sep	Dec	Jan	Jul	Aug	Sep
No. Weekdays*	22	21	23	21	20	23	21
No. Weekend days	9	9	8	10	11	8	9
No. Hours for:							
Parameter 15 †	92	16					
Parameter 16 ‡	16	0	System load data not available in DOE data file				
Parameter 17 §	0	0					
Observed System Load:							
Peak (MW)	1674	1504					
Day/Hour	23/17	8/17	System load data not available in DOE data file				
Load Factor	0.659	0.689					

*Holidays are counted as weekend days.

†Average hourly kWh consumed during all hours of observed system load greater than 80 percent of annual system peak.

‡Average hourly kWh consumed during all hours of observed system load greater than 90 percent of annual system peak.

§Average hourly kWh consumed during all hours of observed system load greater than 95 percent of annual system peak.

The next phase of data reduction was the aggregation of daily usage data to produce weekday and weekend averages. The first two rows of table 3-2 show the numbers of days involved in these aggregations. These values can be utilized to express certain parameter estimates in alternative forms. For

example, although the peak period occurs only on weekdays, parameter 7, average kWh per customer per day during peak period in a selected calendar month, involves an average over all days in the month. Hence, an estimate of parameter 7 can be multiplied by the number of days in the given month and divided by the number of weekdays in the month in order to express the peak-period usage on an average weekday basis.

Following a preliminary analysis of system load data, the hourly consumption data was reduced (aggregated across various days and hours) and edited. These data handling procedures are described in Appendixes A and B. Editing of consumption data involves two steps (1) an automated edit, which excludes customers with large amounts of missing data and flags customers having unusual consumption values (e.g., a high hourly consumption value, or a large number of hours with zero consumption), and (2) a "manual" edit, which allows the user, after examining the consumption patterns of the flagged customers, to delete specific customers from the data base (on a month-by-month basis).

Table 3-3 indicates the number of customers flagged by the automated edit. Hourly usage values for the entire month were manually examined for each of the "flagged" customers. For example, one customer showed zero consumption during the hours of 4:00 P.M. to 8:00 P.M. for 14 consecutive days. Another showed zero usage from midnight to 9:00 A.M. for every day in one month. The predominant pattern that prevailed, however, was the existence of large blocks of consecutive hours of zero consumption. One customer, for example, showed zero values in every hour for the months of July, August, and September.

Table 3-3. Number of Customers Flagged by Month by Test Group
in the Automated Edit

Test Group	1976			1977			
	Aug	Sep	Dec	Jan	Jul	Aug	Sep
Control	0	0	2	3	1	2	2
Experimental	5	2	9	11	8	6	11

Various reasons can be hypothesized for regarding the zero data values as valid. On the other hand, arguments can be made that they are erroneously coded consumption values. Without additional information it is simply impossible to establish for many of the customers whether the zero values are legitimate. For those customers with large blocks of consecutive hours of zero consumption, a decision was made to delete the customer's data for a month if ten or more consecutive days were observed to have zero consumption. If the period of zero usage covered two months, then the customer was deleted for both months. A listing of the customers deleted by this procedure is given in Appendix C.

Table 3-4 shows the results of the automated edit and the subsequent manual edit. The table shows, by month and by test group, the number of customers failing the automated edit (based on the number of missing values) and, hence, excluded from all analyses for that month, as well as those excluded by the manual edit. The number of customers providing valid usage data by treatment and month is given in table 3-5.

Table 3-4. Number of Customers Deleted by Month by Test Group
in the Automated and Manual Edits

Test Group	1976			1977			
	Aug	Sep	Dec	Jan	Jul	Aug	Sep
Control							
Automated Edit	5	5	4	2	5	5	5
Manual Edit	0	0	2	2	1	1	1
Total	5	5	6	4	6	6	6
Experimental							
Automated Edit	9	9	1	2	2	2	2
Manual Edit	1	0	4	5	4	2	5
Total	10	9	5	7	6	4	7

All data reduction and editing procedures are described in Appendixes A and B.

Table 3-5. Customers Providing Usage Data for Analysis

Test Group	1976			1977			
	Aug	Sep	Dec	Jan	Jul	Aug	Sep
Control	50*	50	49	51	49	49	49
Experimental	87*	88	92	90	91	93	90

*The DOE data files provided usage data on only 55 control customers and 97 experimental customers. Usage data on the other 5 control and 3 experimental customers were not available.

3.4 Comparative Analysis Results

For each month selected for analysis, a set of summary statistics were developed which estimated those population parameters pertinent to the evaluation of the effects of a TOU rate. The estimates, coupled with distributional assumptions, were used to test whether significant differences existed between the true population parameters for the two groups. Results of tests of hypotheses concerning these parameters were also developed. The methodology for conducting these tests is described in Appendix B. Appendix D presents a detailed comparison of these data for experimental and control groups in tabular form for August, September, and December 1976, and January, July-September 1977.

Table 3-6 summarizes tests of treatment differences for all months analyzed and identifies parameters for which the experimental TOU rate response differed significantly from the control group response. Entries on this table are shown as percentage changes for those parameters which were considered significant by tests of hypotheses. Blank spaces indicate that the associated differences were not statistically significant at the .05 level. Several patterns are evident in table 3-6. For example, some parameters show significant differences in each month for which data are available, whereas for other parameters, significant differences are not evident in any month. These patterns may be summarized as follows:

Total Consumption. Tests of hypotheses indicated that total monthly consumption (parameter 12) and total consumption for the day of system peak (parameter 6) for the experimental group did not differ significantly from that for the control group. This resulted from the fact that a decrease during the peak period was offset by an increase during the base period even though the latter was not statistically significant. The estimated monthly

Table 3-6. Differences between Control and Experimental Groups for Values Deemed Significant*
(In estimated percentage changes)

Parameter†	1976			1977			
	Aug	Sep	Dec	Jan	Jul	Aug	Sep
Peak Day kWh							
Peak - 1	-39.7	-26.0	--	--	--	--	--
Intermediate - 2	--	--	--	--	--	--	--
Base - 3			--	--	--	--	--
1-hr. before peak period - 4			--	--	--	--	--
1-hr. after peak period - 5			--	--	--	--	--
Total - 6			--	--	--	--	--
Average Day kWh							
Peak - 7	-38.0	-33.2			-27.9	-23.6	-27.6
Intermediate - 8	--	--	--	--	--	--	--
Base - 9							
1-hr. before peak period - 10							
1-hr. after peak period - 11							
Total - 12							
Diversified Demand - 13	-40.2	-29.4	--	--	--	--	--
Noncoincident Maximum Demand - 14							
Average Hourly kWh - 80% - 15	-19.6	-22.4	--	--	--	--	--
Average Hourly kWh - 90% - 16	-34.4	--	--	--	--	--	--
Average Hourly kWh - 95% - 17	--	--	--	--	--	--	--
#1/#6 - 18	-37.0	-25.2	--	--	--	--	--
#7/#12 - 19	-35.4	-29.9	-26.1	-25.6	-24.1	-23.6	-23.9
Class Load Factor (#12/#13 ÷ 24) - 20	60.9		--	--	--	--	--

*Calculated as: $100(\bar{Y}_e - \bar{Y}_c) / \bar{Y}_c$ where $\bar{Y}_e$ and $\bar{Y}_c$ are means for experimental and control groups, respectively. Values are shown only for those cases deemed significant at the .05 level.

†See table A-4 for more general parameter definitions.

-- Indicates zero data.

consumption was slightly higher for the experimental group in three of the seven months examined (December 1976, January and August 1977) and slightly lower in the remaining four months. Tables D-2 through D-8 provide estimates on parameters 6 and 12 for the seven months of the study.

Peak Period. In five of the seven months examined, the experimental group consumed significantly less electricity during the peak period (parameter 7) than the control group. The reduction in consumption ranged from 24 percent in August 1977 to 38 percent in August 1976. December 1976 and January 1977, also showed reductions in usage of approximately 21 percent. However, the differences were not statistically significant because the estimated standard errors for both the control and experimental groups were almost double those observed in the other months. This is probably attributable to high usage by the electric heating customers during these two winter months. Numerically, the difference in consumption for the two groups over the eight months ranged from -1.52 to -2.62 per customer per day (parameter 7 in tables D-2 through D-8). Significant reductions in peak-period usage were also observed for the days of monthly system peak (parameter 1) during August 1976 and September 1976 (the only months for which system data were available).

Usage during the peak period is also reflected through parameters 18 and 19 which measure, respectively, the proportion of total usage in the peak period for the day of system peak and for the month. Since the TOU rate schedule caused a reduction in usage during the peak period but did not affect total consumption, these proportions were significantly lower for the experimental group in each of the seven months.

Base Period. In six of the seven months examined, the experimental customers used more electricity during the base period than the control customers. These increases tend to offset the reduction in consumption

observed in the peak period, suggesting that the TOU rate schedule is causing customers to shift usage from the peak to the base period. The winter months of December and January showed a more pronounced shift in usage from peak to base period. Tests of hypotheses failed to detect the base-period increases as being statistically significant; however, this is understandable. Relative to the peak period, the customer-to-customer variation in usage is much larger during the base period as evidenced by the standard errors for parameters 7 and 9 given in tables D-2 through D-8. Whereas reductions of 2 to 3 kWh per customer per day during the peak period were found to be statistically significant, increases of these magnitudes during the base period cannot be considered statistically significant with the present sample sizes. In fact, because of the increased variability in usage during the base period, a difference (between control and experimental response) would have to range from 5 to 10 kWh per customer per day to be regarded as statistically significant.

Usage during the Hour preceding/following the Peak Period. The variables involving the one-hour periods preceding and following the peak period (parameters 4, 5, 10, and 11) were designed to detect needle peaking for rate structures having two daily rating periods. In general, hourly usage estimates during these periods were higher for customers under the experimental rate (increases were observed in six of the seven months examined--see tables D-2 through D-8). It should be noted, however, that the magnitude of the difference in any one month was not large enough to be considered statistically significant.

Diversified Demand at Time of System Peak and Associated Load Factor. For the two months in which system load data were available (August and September 1976), the diversified demand at the time of system peak (parameter

13) for the experimental group was significantly less (reduction of 0.84 in August and 0.45 in September) than that for the control group. The associated class load factor (parameter 20) for the experimental group was significantly higher in the month of August (0.997 as compared to 0.619). In September, the estimated class load factor for the experimental group was 0.977 as compared to 0.723 for the control group. This difference is not significant at the .05 level but is significant at the .10 level.

Noncoincident Maximum Demand. Noncoincident maximum demand (parameter 14) is derived as the average of the individual customer's maximum demands over the month. Analysis for each of the seven calendar months indicated that noncoincident maximum demand for customers under the experimental rate did not differ significantly from that for customers under the control rate.

In general, it is apparent from tables D-2, D-3, D-7 and D-8 that differences between the experimental and control groups observed in August and September 1976 are essentially the same as those observed in the corresponding months of 1977. This means that the effects of time-of-use pricing of electricity did not change over time (summer 1976 to summer 1977).

3.5 Average Daily Load Curves

Another element in the comparative analysis is the examination of graphic displays of average daily residential load curves. A set of four daily curves were constructed for each of the eight calendar months used in the comparative analysis described earlier in this chapter. Table 3-7 describes the daily load curves constructed for each month. In constructing these load curves, sample results were inflated to represent an average member of the effective population.

Note that a direct relationship exists between the parameters of table 3-1 and the load curves of table 3-7. For example, the summation of usage under load curve 1, day of monthly system peak, in those hours when the peak

price is in effect corresponds to parameter 1. Under a rate structure that does not vary from day to day, a similar summation of usage under load curve 4 corresponds to parameter 7. For these two examples, extending the summation over 24 hours produces parameters 6 and 12, respectively.

Table 3-7. Average Daily Load Curves

Type	Averaging Period
1	Day of monthly system peak
2	Weekend days during month -- holidays included as weekend days
3	Weekdays during the month -- holidays included as weekend days
4	All days during the month

The entire set of load curves for the seven months selected for the study have been included in this chapter. Figure 3-1 shows one set for August 1976 while figures 1-1 and 1-2 in chapter 1 show two of the load curves--average weekday and day of system peak--for the same month in greater detail. Figures 3-1 through 3-7 provide corresponding load curves for all seven months.

The load patterns given in figures 3-1 through 3-7 clearly show the reduction in electricity consumption by the experimental customers during the peak period. Evidence regarding shifting of usage from the peak to the base period is more pronounced in the plots for the winter months of December and January. In both of these months the experimental customers used, on the average, almost 5 kWh per customer per day more during the base period than the control customers. This increase was not statistically significant because of the large variability in usage observed during these months. For the remaining months, the shifts reflected by the small increases in usage during the base period are much less discernible.

Figure 3-1. Average Daily Load Curve for Control and Experimental Rate Schedule, August 1976

Figure 3-2. Average Daily Load Curve for Control and Experimental Rate Schedule, September 1976

35

DAY OF SYSTEM PEAK

ALL DAYS

WEEK DAYS

WEEKEND DAYS

Figure 3-3. Average Daily Load Curve for Control and Experimental Rate Schedule, December 1976

Figure 3-5. Average Daily Load Curve for Control and Experimental Rate Schedule, July 1977

Figure 3-6. Average Daily Load Curve for Control and Experimental Rate Schedule, August 1977

63

ALL DAYS

WEEK DAYS

WEEKEND DAYS

Figure 3-7. Average Daily Load Curve for Control and Experimental Rate Schedule, September 1977

REFERENCES

Dayton Power and Light Company, Method of Selection of Sample Residential Customers to Participate in the Federal Energy Administration Demonstration Project, June 1975.

Ohio Electric Demonstration Project, Product Users Guide, Report prepared by project staff under Cooperative Agreement No. CA-04-50070-00.

Public Utilities Commission of Ohio, Demand Managment Demonstration Project Stage I, Development of Residential Load Characteristics and Stage 4, Demonstration of Residential Incremental Cost Pricing Implemented by Time-of-Day Metering, Final report prepared in cooperation with the Department of Energy, Office of Utility Systems, January 1979.

THIS PAGE
WAS INTENTIONALLY
LEFT BLANK

APPENDIX A: DESCRIPTION OF DATA PROCESSING AND ANALYSIS METHOD

A.1 Analytical Objectives

The proposed data analyses* for the individual rate demonstration projects are described below. Not all of these objectives are relevant to every project.

- 1) Estimation, by treatment, of average daily load curves for the residential class of customers over several different time frames within each of several months (e.g., the day of system peak, weekdays, weekend days).
- 2) Estimation, by treatment, of the average daily residential-class load curve for that day in a month in which the class peak is estimated to occur.
- 3) Estimation of certain parameters characterizing the load curves indicated in (1) and (2).
- 4) Treatment-versus-control comparisons between certain parameters characterizing the load curves indicated in (1).
- 5) Estimation of response surface models which relate customer's usage patterns to electricity prices (perhaps adjusted for certain demographic characteristics).

This appendix provides a brief description of the computer programs which have been written to process, edit, and analyze the data in accordance with the above analytical objectives.

Section A.2 provides an overview of the analysis process, describing the steps involved, their interrelationships, and the purposes of individual programs. Programs and routines used in this analysis were developed at the Research Triangle Institute. For more detailed information see the User's Guide for Processing and Analyzing Data from the Electric Utility Demonstration Projects, revised June 1978. Details on the data editing procedures are provided in section A.3.

* See Analytical Master Plan for the Analysis of the Data from the Electric Utility Rate Demonstration Projects, Volume I, RTI Report No. 41U-1542/00-05F.

A.2. Overview of the Data Processing and Analysis System

As figure A-1 indicates, each analysis begins with an extraction of data for a particular project from the DOE data base. The programs are structured so that only one or two such extractions normally are necessary. For example, the extracted data may cover time periods longer than would be used in any particular analysis, or may include data on customers assigned to treatments which are not relevant to a particular analysis. Selections of treatments and months can be made at any stage in the analysis process. Figure A-1 indicates the structure of the analysis process, including the intermediate data files necessary for producing the final product. The five analytical objectives are satisfied as follows:

- 1) Plot of data from data file U-4,
- 2) Plot of (selected) data from data file U-6,
- 3) Output from the comparative analysis program and the daily load curve and system load curve estimation programs,
- 4) Output from the comparative analysis and/or the analysis of covariance program,
- 5) Output from the response surface analysis program.

It would be desirable to build these programs into an integrated system to minimize the user's input requirements and reduce the number of individual program runs. This does not appear to be feasible, however, because examination of the output of one program is generally necessary before execution of a subsequent program. In addition, the general system may have to be supplemented by other special-purpose programs (or enhancements to the general programs which can be implemented on an optional basis) on a project-specific basis. Three areas where such special programs are likely to be required are the following:

Figure A-1. Overview of Analysis System

NOTE: Dots indicate that a sorting of the data file may be required.

- 1) additional types of manual editing,
- 2) recoding of strata codes (e.g., in order to collapse certain strata into a single stratum), and
- 3) aggregation across monthly data records in order to perform analyses on a seasonal basis.

The remainder of this section briefly describes the individual programs and their inputs and outputs.*

A.2.1 System Load Analysis

This program extracts system load data for that day in each calendar month during which the hourly system peak for the month occurred. A data file (SL-1) containing this information is written. A second data file (SL-2) is also produced; this file simply represents a reformatting of the SL-0 file. The program output can be used to select those months for which subsequent analyses should be performed. In addition, if sufficient system load data are available, definition of the annual system peak is possible. The basic input to the program is the hourly system load data (SL-0), one record per day, in the form generated by a DAY BY HOUR extraction using the DOE extraction system.

The program writes a monthly record for each month having a "sufficient" number of valid hourly system load readings. "Sufficient" is defined by the program parameter SLIM, which gives the proportion of hours required in the month. The current value of SLIM is .95, indicating that 95 percent or more of the hourly system load readings must be positive, that is, non-missing.

* The response surface analysis and the (optional) analysis of covariance are handled by packaged routines and are not included in the discussion.

A.2.2 Linkage of Treatment, Stratum, and Weather Station Codes to U-0

This program augments treatment, stratum, and weather station codes from the DEMO-0 data file onto the usage data file U-0 by matching customer ID numbers from the two files. The resulting file, U-1, like U-0, contains one record per customer per day. Printed output consists of a list of customer ID numbers for which no match was achieved.

A.2.3 Automated Edit, Data Reduction

This program has four major purposes. It creates a reduced set of usage data, performs an automated editing of the hourly usage data, and provides outputs useful for performing additional editing. In addition, it provides counts of respondents, by treatment and stratum. The basic data reduction involves converting a data file (U-1) containing one record per customer per day to a data file containing from 3 to 7 records per customer per (calendar) month. Average hourly usage per customer is constructed over seven different time-frames (corresponding to the seven record types) which are as follows:

<u>Type</u>	<u>Time Frame</u>
1	day of monthly system peak
2	weekdays (excluding holidays) in calendar month
3	weekend days (including holidays) in calendar month
4	all days in calendar month in which system load exceeds $Q_1 \times P$
5	all days in calendar month in which system load exceeds $Q_2 \times P$
6	all days in calendar month in which system load exceeds $Q_3 \times P$
7	all days in calendar month

where P is a user-supplied value of the annual system peak and the Q_i are program parameters (currently set at $Q_1 = .80$; $Q_2 = .90$, $Q_3 = .95$).

The program writes usage records only for those customer-months with sufficient non-missing usage data in U-1. Thus U-2 represents an edited file. In addition, since some zero usage values may actually be miscoded missing values, the program performs a re-edit of the U-1 data. This involves performing the edit under the assumption that all zero kWh readings represent missing values. Records in U-2 which fail this re-editing procedure are flagged. Records derived from usage data in U-1 that contain one or more suspiciously-large hourly kWh readings during the month are also flagged.

The program requires four types of data (if all record types are to be built). A file of usage data (U-1) and the user's control data are always required, while the system load data files SL-1 and SL-2 may be required in some cases.

Output consists of two data files, in addition to printed output. One file is the index file I-1, which summarizes each customer's status with respect to the editing performed by the program. This file contains one record per customer per (selected) month -- for each customer having usage data on file U-1 in the particular month. The second output file is U-2, which contains the reduced, edited usage data. Tables A-1 and A-2 show, respectively, the information contained in each record of the I-1 and the U-2 data files.

Two major types of printed output occur. The first corresponds closely to the information contained in the index file I-1. For customers with a flag value of 2 in a particular month, the pattern of the hourly usage data during the month is displayed. This output can be used as an aid for determining whether or not the customer's usage data should be included in the analysis. The second major portion of the

TABLE A-1. Form of Records in File I-1

<u>Variable</u>	<u>Description</u>
JDCUST	Customer number
JMONTH	Year/Month
JTRT	Treatment code
JSTRAT	Stratum code
IFLAG	Flag indicating customer status (0 = no usage records written on U-2 due to edit failure 1 = some usage records written on U-2; all records written on U-2 pass both edit and re-edit 2 = some usage records written on U-2; at least one record failed to pass the re-edit)
Z2	Maximum hourly usage occurring on weekdays (excluding holidays) in month
Z3	Maximum hourly usage occurring on weekends (including holidays) in month
MB(I) (I=1,2,...,7)	No. of missing hourly values during time frame (record type) I
KB(I) (I=1,2,...,7)	No. of missing or zero hourly values during time frame (record type) I
IVEC(1)	Edit indicator for record type 1
IVEC(2)	Edit indicator for record type 2, 3, and 7
IVEC(3)	Edit indicator for record type 4
IVEC(4)	Edit indicator for record type 5
IVEC(5)	Edit indicator for record type 6
JVEC(1)	Re-edit indicator for record type 1
JVEC(2)	Re-edit indicator for record type 2, 3, and 7
JVEC(3)	Re-edit indicator for record type 4
JVEC(4)	Re-edit indicator for record type 5
JVEC(5)	Re-edit indicator for record type 6

NOTE: IVEC(I) (JVEC(I)) is equal to 1 if edit (re-edit) is passed, is equal to 0, otherwise.

TABLE A-2. Form of Records in File U-2

<u>Variable</u>	<u>Description</u>
JDCUST	Customer number
JW	Weather station code
JMONTH	Year/Month
JTRT	Treatment code
JSTRAT	Stratum code
J	Record type (Time Frame Indicator)
IFLAG	Flag indicating customer status (see Table A-1)
YM(J)	Vector of average hourly kWh values -- over time frame J
MM(J)	Vector of sample sizes (no. of non-missing values) used for computing YM(J) values
KM(J)	Vector containing the no. of positive values entering into YM(J) values
Z(J)	For J=1,2,3,7: maximum hourly kWh during time frame J; For J=4,5,6: average hourly kWh over all hours in month for which system load exceeds $Q_{J-3} \times P$
ND(J)	For J=1: the day of system peak; Otherwise: the number of potential* days in time frame J over which YM(J) values are computed
NH(J)	For J=1: the hour of system peak; Otherwise: the number of potential* hours over which Z(J) values are computed
MB(J)	No. of hours with non-missing kWh values involved in the computation of Z(J) values
KB(J)	No. of positive kWh values involved in the computation of Z(J) values

* That is, if no missing values are present.

output from this program is an array of sample sizes, by month, treatment, stratum, and record type. This can be used to determine where problems of estimability may arise in subsequent analysis programs.

A.2.4 Manual Edit

This program allows the user to perform additional editing of the U-2 data file. These edits simply involve flagging particular records (which are identified through the user's control data). Flagging of a record with record type equal to 2, 3, or 7 implies flagging of all existing record types for that customer and month. This rule is consistent with the rules employed in the automated edit. Since all subsequent programs utilizing U-3, the updated version of U-2, skip these flagged records, the update effectively corresponds to a deletion of the indicated records. The editing is documented in the index file I-2, which is the updated version of I-1.

A.2.5 Load Curve Estimation

This program estimates average daily load curves for the residential population (or a specified subpopulation thereof), where the time averaging corresponds to the seven time frames described in Section A.2.3. The program simply computes a weighted sum, by treatment and month, of the individual usage records in data file U-3. Within each stratum, the customers with data in U-3 are assumed to constitute a random sample from the effective population. Strata sizes are thus required as part of the user's input. Output can be restricted to particular months, treatments, and record types. A data file containing the estimated average daily load curve values is generated (file U-4 in Figure A.1). Printed output consists of a list of the strata and sample size information and (optionally) of the estimated load values.

A.2.6 Creation of Analysis Variables

This program creates a file of analysis variables for the comparative analyses (or certain dependent variables for the response surface analyses) from the data file U-3. The output file, U-5, contains one record per customer per (selected) month; each record contains values of 17 usage variables, as shown in Table A-3. Construction of many of these variables requires a definition of a pricing scheme that indicates those seasons, times-of-week, and hours within days during which peak-, intermediate-, and base-level prices are in effect. This pricing scheme is defined through the user's input; only a single pricing scheme is permitted.

Analysis variables Z1-Z6 and Z13 are constructed from input record type 1. The variables Z1-Z5 are constructed in accordance with the pricing scheme supplied by the user's control data.

Analysis variables Z15, Z16, Z17, and Z14 are extracted directly from the Z(J) variables appearing, respectively, in record types 4, 5, 6, and 7 of data file U-3 (see Table A-2).

Analysis variable Z12 is computed as the sum of the elements in the YM(7) vector (on record type 7 of data file U-3). Analysis variables Z7-Z11 are also derived from record type 7, if a single weekly rating period is indicated for the month. If two weekly rating periods are indicated, these variables are constructed from record types 2 and 3. The weekday pricing scheme is then associated with record type 2 and the weekend scheme, with record type 3.

A.2.7 Comparative Analysis

The purpose of this program is to estimate population parameters, by treatment and month, and to make selected comparisons of

TABLE A-3. Form of the U-5 Data File

<u>Variable</u>	<u>Description</u>
IDCUST	Customer number
IW	Weather station code
JMONTH	Year/Month
KTRT	Treatment code
KSTRAT	Stratum code
IFLAG	Editing flag
Z1	kWh during peak period on day of observed system peak in JMONTH
Z2	kWh during intermediate period on day of observed system peak in JMONTH
Z3	kWh during base period on day of observed system peak in JMONTH
Z4	kWh during one-hour periods preceding peak periods--on day of observed system peak in JMONTH
Z5	kWh during one-hour periods following peak periods--on day of observed system peak in JMONTH
Z6	kWh on day of observed system peak in JMONTH
Z7	Average daily peak-period usage in JMONTH
Z8	Average daily intermediate-period usage in JMONTH
Z9	Average daily base-period usage in JMONTH
Z10	Average daily usage for JMONTH during one-hour periods preceding peak periods
Z11	Average daily usage for JMONTH during one-hour periods following peak periods
Z12	Average daily usage for JMONTH
Z13	Usage during one-hour corresponding to hour of observed system peak in JMONTH
Z14	Maximum hourly usage during JMONTH
Z15	Average hourly usage during hours in which observed system load exceeded .80 times the (assumed) annual system peak
Z16	Same as Z15 with .90 replacing .80
Z17	Same as Z15 with .95 replacing .80

NOTE: Input records (file U-3) with IFLAG values other than 1 or 2 are skipped; hence, only values of 1 and 2 will occur for IFLAG on the output file U-5.

NOTE: Variables Z7, Z8, Z9, Z10 and Z11 are averaged over all days in the month even though peak, intermediate, and base prices may be present only during a subset of the days. Holidays are treated as defined by the user in the creation of the U-2 data file. It is assumed, however, that the monthly system peak does not occur on a weekday which is one of user-defined holidays, since input to this program does not allow it to identify that such a day should be treated like a weekend day in terms of the pricing scheme.

NOTE: Any of the variables Z1 through Z17 which cannot be constructed due to lack of the necessary input records are coded as -1.

the treatments with respect to these population parameters. The program assumes that customers with data in file U-5 constitute random samples from each stratum of the effective population. Strata and sample size information -- furnished by the user -- are used to compute the weighted estimates. The user's control data are used to restrict the estimates to particular months and treatments. Comparisons are made between treatment 1 and each of the other treatments, where the order of the treatments is specified by the user. Generally, treatment 1 would be a control group, although this is not required to be the case. The estimates involved in the comparisons are assumed to be approximately normally distributed; an approximate t-test is used for testing the equality of the various treatment population parameters.

The output consists of a series of (NTRT-1) tables for each month. The first table presents results for treatments 1 and 2; the second, for treatments 1 and 3, etc. For each of the two treatments involved in a particular table, the following are given for each parameter of interest:

- a) sample size upon which the parameter estimate is based,
- b) the value of the parameter estimate, and
- c) the estimated standard error of the estimate.

An estimate is not provided for the particular treatment and parameter if there is any stratum which contains no customers (who furnish data) assigned to that treatment. Similarly, item (c) is not given if any stratum contains no customers or a single customer on the particular treatment.

If all three of the above quantities can be determined for each of the two treatments, a test of significance is performed. The following output is presented:

- a) the estimated difference between the population parameters under treatment 1 and treatment k ($k = 2, 3, \dots, \text{NTRT}$), and an indication of the statistical significance of the true difference *
- b) the standard error of this difference
- c) approximate 99 percent confidence limits for the true difference.*

The above estimates and comparisons are made for the 20 parameters given in table A-4. The parameters are labeled numerically in the output as indicated in this table.

A.2.8 Daily Load Curve Estimation

This program provides average hourly usage estimates (per customer) for the specified (residential) population. These estimates are provided for each selected treatment at each hour of each day in each selected month. In addition, the estimated maximum monthly diversified demand (per customer) for this class of customers is determined for each treatment, as well as its time of occurrence. The monthly class load factor is also presented. Customers whose data are to be utilized in making the estimates are determined by the flag and IVEC values appearing on the index file I-2. The data file U-1 is the source file for the usage data. The estimated residential load curves are written onto data file U-6, which can subsequently be used for producing graphical output.

A.2.9 Daily System Load Curve Estimation

This program combines the existing daily system load curves (data file SL-2) with the daily residential load curves estimated for each treatment (data file U-6) to produce estimated system load

* The tests and confidence limits assume that degrees of freedom for the test statistic (computed as a t statistic) are sufficiently large so that comparisons to tabulated Normal (0,1) percentage points are justified.

curves for each treatment. The user must furnish control data in addition to these two data files. The estimation procedure assumes that all pertinent system characteristics (e.g., distribution and transmission losses) can be reduced to a single user-supplied constant which holds uniformly over both time and treatments. Output consists of a list of the estimated daily system load curves, by treatment. The following system-level quantities are also presented for each selected month and treatment.

- average hourly system load
- maximum hourly system load
- system load factor
- day and hour at which maximum hourly system load occurs

A.3 Data Editing Procedures

Let the system load data on file SL-2 be defined as

S_{hjm} = system load at hour h of j^{th} day in m^{th} calendar month.

Let $(h^*, j^*)_m$ denote the hour and day in the m^{th} month at which the maximum hourly load, S_m^* , occurs. Let A_m be the (user-supplied) value of the (assumed) annual system peak. Now define

$$G_{hjm\ell} = \begin{cases} 1 & \text{if } S_{hjm} > C_{\ell} A_m \\ 0 & \text{otherwise,} \end{cases}$$

where C_{ℓ} are program parameters with current values as follows:

ℓ	C_{ℓ}
1	.80
2	.90
3	.95

Note that the data S_{hj^*m} is contained on the data file SL-1. We also define the following quantities:

TABLE A-4. List of Parameters Estimated in the Comparative Analysis (By Treatment and Month)

1. Average kWh per customer during peak period on day of observed system peak in calendar month i
2. Average kWh per customer during intermediate period on day of observed system peak in calendar month i
3. Average kWh per customer during base period on day of observed system peak in calendar month i
4. Average kWh per customer during 1-hr. preceding peak period(s) on day of observed system peak in calendar month i
5. Average kWh per customer during 1-hr. following peak period(s) on day of observed system peak in calendar month i
6. Average kWh per customer on day of observed system peak in calendar month i
7. Average kWh per customer per day during peak period for i th calendar month
8. Average kWh per customer per day during intermediate period for i th calendar month
9. Average kWh per customer per day during base period for i th calendar month
10. Average kWh per customer per day during 1-hr. preceding peak period(s) for i th calendar month
11. Average kWh per customer per day during 1-hr. following peak period(s) for i th calendar month
12. Average kWh per customer per day during i th calendar month
13. Diversified demand (average per customer) at time of observed system peak in the i th calendar month
14. Noncoincident maximum demand (average per customer) for i th calendar month
15. kWh per customer averaged over all hours in the i th calendar month for which observed system load was greater than .80 of observed annual system peak
16. kWh per customer averaged over all hours in the i th calendar month for which observed system load was greater than .90 of observed annual system peak
17. kWh per customer averaged over all hours in the i th calendar month for which observed system load was greater than .95 of observed annual system peak
18. Proportion of total kWh in peak period on day of observed system peak in the i th calendar month ($\#1/\#6$)
19. Proportion of total kWh in peak period for i th calendar month ($\#7/\#12$)
20. Class load factor for i th calendar month based on diversified demand at time of observed system peak ($\#12/\#13 \div 24$)

$$X_{jm}^{(1)} = \begin{cases} 1 & \text{if day } j \text{ is the day of system peak in month } m \\ 0 & \text{otherwise} \end{cases}$$

$$X_{jm}^{(2)} = \begin{cases} 1 & \text{if day } j \text{ of month } m \text{ is a weekday (excluding holidays)} \\ 0 & \text{otherwise} \end{cases}$$

$$X_{jm}^{(3)} = 1 - X_{jm}^{(2)}$$

$$X_{jm}^{(4)} = \max_h [G_{hjm1}]$$

$$X_{jm}^{(5)} = \max_h [G_{hjm2}]$$

$$X_{jm}^{(6)} = \max_h [G_{hjm3}]$$

The above indicator variables define subsets of days in month m corresponding to the first six output record types. Let $Y_{hjm}(i)$ denote the usage for the i^{th} customer at hour h of day j in month m . Define

$$M_{hjm}(i) = \begin{cases} 1 & \text{if } Y_{hjm}(i) \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{i.e., if } Y_{hjm}(i) \text{ has a non-missing value})$$

and

$$K_{hjm}(i) = \begin{cases} 1 & \text{if } Y_{hjm}(i) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The average usage vector for the i^{th} customer over time frame J in month m is then computed as

$$\bar{Y}_{h,m}^{(J)}(i) = \frac{\sum_{j=1}^{r_m} X_{jm}^{(J)} Y_{hjm}(i) K_{hjm}(i)}{\sum_{j=1}^{r_m} X_{jm}^{(J)} M_{hjm}(i)}$$

where r_m is the total number of days in month m . The vectors $YM(J)$, $MM(J)$, and $KM(J)$ of Table A-2 correspond, respectively, to

$$\bar{Y}_{h.m}^{(J)}(i) \quad h = 1, 2, \dots, 24;$$

$$\sum_{j=1}^{r_m} X_{jm}^{(J)} M_{hjm}(i) \quad h = 1, 2, \dots, 24;$$

and

$$\sum_{j=1}^{r_m} X_{jm}^{(J)} K_{hjm}(i) \quad h = 1, 2, \dots, 24.$$

The average vector for record type 7 is computed as a weighted average of the weekend and weekday usage vectors:

$$\bar{Y}_{h.m}^{(7)}(i) = \left(X_{.m}^{(2)} \bar{Y}_{h.m}^{(2)} + X_{.m}^{(3)} \bar{Y}_{h.m}^{(3)} \right) / r_m$$

where

$$X_{.m}^{(J)} = \sum_{j=1}^{r_m} X_{jm}^{(J)}.$$

Note that this calculation involves an (implicit) imputation of the raw usage data if missing values occur.

The $Z(J)$ variables (see table A-3) are computed as follows:

$$Z_m^{(J)}(i) = \max_{h,j} \left[X_{jm}^{(J)} Y_{hjm}(i) \right] \quad \text{for } J=1,2,3;$$

$$Z_m^{(7)}(i) = \max \left[Z_m^{(2)}(i), Z_m^{(3)}(i) \right];$$

and

$$Z_m^{(J)}(i) = \sum_{j=1}^{r_m} \sum_{h=1}^{24} G_{hjm, J-3} Y_{hjm}^{(i)} K_{hjm}^{(i)} / R_m^{(J)}(i)$$

for $J=4,5,6$, where

$$R_m^{(J)}(i) = \sum_{j=1}^{r_m} \sum_{h=1}^{24} G_{hjm, J-3} M_{hjm}^{(i)}.$$

Output records are written only if "sufficient" non-missing data is available in the U-1 file. In order to make this determination, the program compares the number of non-missing values to the potential number of hourly values which are involved in the calculations of the $\bar{Y}_{h.m}^{(J)}(i)$ and the $Z_m^{(J)}(i)$ variables. In particular, three basic types of screening conditions are employed:

$(a) \min_h \left[M_{h.m}^{(J)}(i) / X_{.m}^{(J)} \right] < .75$	$J=1,2,\dots,6$	(i.e., data is available for at least 75% of the potential number of days -- at <u>each</u> hour of the day)
$(b) \sum_{h=1}^{24} M_{h.m}^{(J)}(i) / 24 X_{.m}^{(J)} < .90$	$J=1,2,\dots,6$	(i.e., data is available for at least 90% of the potential number of hours
$(c) R_m^{(J)}(i) / G_{..m, J-3} < .90$	$J=4,5,6$	

where

$$G_{..m, J-3} = \sum_{j=1}^{r_m} \sum_{h=1}^{24} G_{hjm, J-3}$$

Note that $X_{.m}^{(J)}$ represents the potential number of hourly observations over which averages of the Y-type variables can be constructed for record type J in month m; $G_{..m, J-3}$ is the potential number of hours for averages of the Z-type variables.

The values appearing on the right-hand side of the inequalities are program parameters. The above screening criteria are logically ordered as follows:

1. No records for customer i in month m are retained if condition (a) holds when J=2 or 3 or condition (b) holds when J=2 or 3. Otherwise record types 2, 3, and 7 are retained and the editing continues.
2. For J=1, 4, 5, and 6, record type J is not retained if either (a) or (b) holds for the corresponding value of J.
3. Record type J, J=4, 5, 6, is not retained if condition (c) holds for the corresponding value of J.

It should be noted that all of the edit checks performed by the program are concerned with the amount of missing data. No checks involving the patterns of missing data are made.

Because zero-usage hourly data may actually represent miscoded missing values, and because extremely large usage values may represent invalid readings, additional checks are made and reported in the output. These involve performing a re-edit of the hourly usage data under the assumption that all zero-kWh readings actually represent missing values. If any (output) record type passed the editing procedure but fails to do so under this assumption, this customer's records are flagged with a value of 2. A flag value of 2 also is assigned when the maximum hourly usage value during the month exceeds 20 kWh. For any customer receiving a flag value 2, the printed output shows the customer's usage pattern. This output appears as a day-by-hour array employing the following symbolic notation:

<u>Symbol</u>	<u>Meaning</u>
*	Usage value missing
0	Usage value of zero kWh
(blank)	$0 < \text{usage value} \leq 10 \text{ kWh}$
1	$10 < \text{usage value} \leq 20 \text{ kWh}$
2	$20 < \text{usage value} \leq 30 \text{ kWh}$
:	:
:	:
8	$80 < \text{usage value} \leq 90 \text{ kWh}$
9	Usage value $> 90 \text{ kWh}$.

These printouts are useful for determining if additional records should be deleted prior to analysis via the manual edit. Both the manual and automated edits are aimed at including or excluding entire records, as opposed to specific data elements within records. This approach is advocated for two reasons. First, it facilitates the analysis by applying a common sampling weight to all variables within a record type. For instance, load curve estimates and estimates of parameters characterizing the load curve are based on a common set of sample customers. Secondly, it is likely to produce estimates with less bias in that the presence of large amounts of missing or invalid data suggests that the time intervals associated with the raw usage data may also be in error; hence, data which appears valid may, in fact, be invalid.

As the output records of the U-2 file are produced, a count is made. All counts are reported in the printed output in the form of a table of sample sizes broken down by month, treatment, stratum, and record type. Examination of this table will indicate whether problems of estimability will occur in the analysis due to sample size counts of zero or one. Any invalid (i.e., out of range) stratum codes which may have occurred will appear in this table in stratum number zero.

APPENDIX B: STATISTICAL ESTIMATION AND TESTING PROCEDURES USED IN COMPARATIVE ANALYSIS

B.1 Population and Sampling Structures

The population of customers for which inferences are desired is either the entire class of residential customers in a utility's service area or some specific subgroup of this class. Valid statistical inferences to this population is, in a strict sense, impossible, because the population is constantly changing. Target populations for most of the demonstration projects have been defined (at least, initially) as all households with an active residential account as of a specific time (namely, when the sample was drawn). This appears to be a reasonable modification of the desired population of inference. To obtain valid estimates for this target population requires that a probability sample of the target population be exposed to each experimental rate.

In most of the demonstration projects, however, a number of customers in the target population have been systematically excluded from participation in the experiment. These exclusions are of two basic types.

Type I exclusions involve those customers in the target population who have a zero probability of inclusion in the experiment. Although this type of exclusion may reduce the meaningfulness of results, since the sampled population may be less relevant, it does not otherwise cause inferential problems as long as inferences are limited to this reduced population.

Typical exclusions of this type are the following:

- 1) customers with less than a 12-month billing history,
- 2) customers with an estimated bill in the past X months,

- 3) accounts associated with seasonal dwellings*,
- 4) customers with irregular monthly usage patterns,
- 5) customers with low annual usage or low usage in X out of the past Y months,
- 6) customers with credit problems,
- 7) multiple-family dwellings*, and
- 8) customers living in certain geographic areas within the service area*.

Type II exclusions differ from Type I in that the total number of eligible accounts in the resulting effective population is unknown when exclusions of the second type occur. Type II exclusions may be further subdivided according to whether they occur because of customers' behavior or because of other circumstances. The latter include:

- 1) wiring incompatibility or meter inaccessibility, and
- 2) failure to locate or contact the customer.

Customer behavior which may bias results includes:

- 1) failure to be included because of numerous customer complaints,
- 2) failure to participate in experiments (including, in some cases, failure to volunteer as a participant) or cooperate in interviews, and
- 3) failure to be included because of certain interview responses (e.g., an indication of an intent to move in the near future).

B.2 Notation

Let N^* denote the number of residential customers in the target population and N , the number in the sampled population. Assume that the sampled population is partitioned into H strata, with $N(h)$ customers in

* These are considered as Type I exclusions only if such information is available for all accounts. Otherwise, they are regarded as Type II exclusions. It should also be noted that these may be legitimate exclusions if, in practice, such customers would not be placed on a a time-of-use rate.

the h^{th} stratum ($h = 1, 2, \dots, H$). Because of the type II exclusions, it is apparent that implementation of a time-of-use rate (under the same conditions as used in the experiment) would not expose the entire sampled population to such a rate. Under these conditions, then, one cannot make valid statistical inferences to the entire sampled population, but only to those customers who would have been experimental participants if given the opportunity. Assume that M customers are in this population of potential participants (i.e., in the effective population -- all customers not belonging to either of the exclusion categories); in the h^{th} stratum, assume that there are $M(h)$ customers in the effective population. This population structure is depicted in figure B.1.a.

As indicated in figure B.1.b, the sampling structure within each stratum generally involved the following steps:

- 1) drawing a random sample
- 2) screening a sufficient number of the selected customers so as to achieve a fixed sample size of participants
- 3) assigning these participants at random to each of several electricity rates.

Let K denote the number of rates (treatments) involved in the experiment (including, in general, a control group). Let $m(h)$ denote the number of sample participants in the h^{th} stratum and let $n(h)$ denote the number of customers screened in order to achieve this many participants. At a particular point in time, not all of the customers will provide valid consumption data; out of the $m_k(h)$ participants assigned (at random) to treatment k , assume that $r_k(h)$ do provide valid consumption data. In particular, let $Y(h,i)$ and $Z(h,i)$ denote the values of two unique functions of the consumption data for the i^{th} participant of the h^{th} stratum. Let $T_k(h,i)$ be equal to one if the i^{th} participant of the h^{th}

Figure B-1. Population and Sampling Structures

a. POPULATION
STRUCTURE:

b. SAMPLING STRUCTURE WITHIN
STRATUM (h) OF THE SAMPLED
POPULATION:

stratum is assigned to treatment k, and equal to zero, otherwise. Let $V_y(h,i)$ be equal to 1 if the i^{th} customer of h^{th} stratum provides a valid value of $Y(h,i)$, and equal to 0, otherwise. Note that

$$r_k(h) = \sum_{i=1}^{n(h)} V_y(h,i) T_k(h,i) \quad \text{and} \quad m_k(h) = \sum_{i=1}^{n(h)} T_k(h,i).$$

B.3. Estimation of Averages for the Effective Population

To estimate the average of a consumption variable Y for the effective population (if this population were exposed to the k^{th} treatment) requires that one estimate both the size of the effective population (M) and the total of the variable over the population (for the given treatment). It is assumed that a customer's participation or non-participation does not depend upon the particular treatment to which he has been assigned. The size of effective population can therefore be estimated as

$$\hat{M} = \sum_{h=1}^H \hat{M}(h)$$

where

$$\hat{M}(h) = N(h) \frac{m(h)-1}{n(h)-1}.$$

This estimate assumes that the number of rejections in the h^{th} stratum, $n(h)-m(h)$, represents an observation from a negative binomial distribution with parameters $m(h)$ and $p(h)$, where $p(h)$ is the probability that a customer (in this stratum) will be a participant in the experiment.

The total of the consumption variable Y--over the effective population exposed to treatment k--is estimated as

$$\hat{Y}_k = \sum_{h=1}^H \hat{M}(h) \hat{\bar{Y}}_k(h, \cdot)$$

where

$$\hat{\bar{Y}}_k(h, \cdot) = \sum_{i=1}^{n(h)} v_y(h, i) T_k(h, i) Y(h, i) / r_k(h)$$

is the estimated mean for stratum h when treatment k is in effect. In this formulation it is implicitly assumed that those customers providing valid data in the h^{th} stratum still constitute a random sample from the h^{th} stratum of the effective population.

The mean of Y for the effective population under treatment k is then estimated as the ratio

$$\hat{\bar{Y}}_k = \hat{\bar{Y}}_k / \hat{M}. \quad (B-1)$$

The estimated variance of $\hat{\bar{Y}}_k$ is approximated as

$$\widehat{\text{Var}} \left[\hat{\bar{Y}}_k \right] = \frac{1}{\hat{M}^2} \left[\widehat{\text{Var}}(\hat{\bar{Y}}_k) + \frac{\hat{\bar{Y}}_k^2}{\hat{M}^2} \widehat{\text{Var}}(\hat{M}) - 2 \frac{\hat{\bar{Y}}_k}{\hat{M}} \widehat{\text{Cov}}(\hat{\bar{Y}}_k, \hat{M}) \right]$$

This can be rewritten as

$$\widehat{\text{Var}} \left[\hat{\bar{Y}}_k \right] = \frac{1}{\hat{M}^2} \sum_{h=1}^H \left\{ \left(\left[\hat{M}(h) \right]^2 + \hat{Q}(h) \right) \frac{s_k^2(h)}{r_k(h)} + \left(\hat{\bar{Y}}_k(h, \cdot) - \hat{\bar{Y}}_k \right)^2 \hat{Q}(h) \right\} \quad (B-2)$$

where

$$s_k^2(h) = \sum_{i=1}^{n(h)} v_y(h, i) T_k(h, i) \left[Y(h, i) - \hat{\bar{Y}}_k(h, \cdot) \right]^2 / [r_k(h) - 1]$$

and

$$\hat{Q}(h) = \widehat{\text{Var}} \left[\hat{M}(h) \right] = \hat{M}(h) \left[N(h) - \hat{M}(h) \right] / [n(h) - 2].$$

Note that if $\hat{M}(h) = N(h)$ for all h (i.e., a 100% participation rate), then equation (B-1) produces the estimated mean for the sampled population and, since $\hat{Q}(h) = 0$, equation (B-2) produces the appropriate variance for this estimate.

B.4 Estimation of Ratios of Averages for the Effective Population

Load factors and certain proportions are estimated for the effective population by combined ratio estimates of the form

$$\hat{R}_k = \frac{\hat{Y}_k}{\hat{Z}_k} = \frac{\sum_{h=1}^H \hat{M}(h) \hat{Y}_k(h, \cdot)}{\sum_{h=1}^H \hat{M}(h) \hat{Z}_k(h, \cdot)}.$$

The approximate variance of $\hat{R}_k$ is estimated as

$$\widehat{\text{Var}}(\hat{R}_k) = \frac{1}{\hat{Z}_k^2} \sum \left\{ \left(\left[\hat{M}(h) \right]^2 + \hat{Q}(h) \right) \hat{r}_k(h) + \left[\hat{\Delta}_k(h) \right]^2 \hat{Q}(h) \right\}$$

where

$$\hat{\Delta}_k(h) = \hat{Y}_k(h, \cdot) - \hat{R}_k \hat{Z}_k(h, \cdot),$$

$$\hat{r}_k(h) = \widehat{\text{Var}} \left[\hat{Y}_k(h, \cdot) \right] + \hat{R}_k^2 \widehat{\text{Var}} \left[\hat{Z}_k(h, \cdot) \right]$$

$$- 2\hat{\rho}_k(h) \hat{R}_k \sqrt{\widehat{\text{Var}} \left[\hat{Y}_k(h, \cdot) \right] \widehat{\text{Var}} \left[\hat{Z}_k(h, \cdot) \right]}$$

and $\hat{\rho}_k(h)$ = the estimated correlation between $Y(h, i)$ and $Z(h, i)$ for customers in stratum h and assigned to treatment k .

B.5 Interval Estimates for the Difference in Effective Population Parameters Under Two Different Treatments

Without loss of generality, assume that treatment k ($k=2, 3, \dots, K$) is to be compared to treatment 1 with respect to a consumption variable

Y. More explicitly, assume that one wants a $(1-\alpha) \times 100\%$ confidence interval for the difference $\bar{Y}_k - \bar{Y}_1$. Approximate confidence limits are computed by

$$\hat{\bar{Y}}_k - \hat{\bar{Y}}_1 \pm C_\alpha \sqrt{\widehat{\text{Var}} \left[\hat{\bar{Y}}_k \right] + \widehat{\text{Var}} \left[\hat{\bar{Y}}_1 \right] - 2\hat{\beta}_k} \quad (\text{B-3})$$

where C_α is the $(1 - \frac{\alpha}{2}) \times 100$ percentage point of the standard normal distribution, and

$$\hat{\beta}_k = \frac{1}{\hat{M}^2} \sum_{h=1}^H \hat{Q}(h) \left[\hat{\bar{Y}}_k(h, \cdot) - \hat{\bar{Y}}_k \right] \left[\hat{\bar{Y}}_1(h, \cdot) - \hat{\bar{Y}}_1 \right].$$

The null hypothesis that $\bar{Y}_1 = \bar{Y}_k$ is rejected at the α level of significance if the upper and lower confidence limits (B-3) are of the same sign.

Approximate $(1-\alpha) \times 100\%$ confidence limits for the difference $R_k - R_1$ are given by

$$\hat{R}_k - \hat{R}_1 \pm C_\alpha \sqrt{\widehat{\text{Var}}(\hat{R}_k) + \widehat{\text{Var}}(\hat{R}_1) - 2\hat{\gamma}_k}$$

where

$$\hat{\gamma}_k = \frac{1}{\hat{Z}_k \hat{Z}_1} \sum_{h=1}^H \hat{Q}(h) \hat{\Delta}_k(h) \hat{\Delta}_1(h).$$

Appendix C: Listing of Customers Deleted in Manual Edit

[illegible]

MONTH = 7608 , FLAG = 3

4086	EXP	1	0	123	24	24	0	0	147	4.78	4.80
------	-----	---	---	-----	----	----	---	---	-----	------	------

MONTH = 7612 , FLAG = 3

1044	CNT	1	0	120	11	0	0	0	131	5.29	3.90
1075	CNT	1	0	160	48	0	0	0	208	7.12	6.19
4006	EXP	1	0	213	48	0	0	0	261	4.05	3.72
4016	EXP	1	0	465	192	0	0	0	657	5.41	0.00
4034	EXP	1	0	251	96	0	0	0	347	4.87	3.88
4046	EXP	1	0	157	48	0	0	0	205	3.18	3.20

MONTH = 7701 , FLAG = 3

1044	CNT	1	0	504	240	0	0	0	744	0.00	0.00
1075	CNT	1	0	504	240	0	0	0	744	0.00	0.00
4006	EXP	1	0	504	240	0	0	0	744	0.00	0.00
4011	EXP	1	0	250	96	0	0	0	346	3.73	5.49
4020	EXP	1	0	302	144	0	0	0	446	13.00	12.13
4046	EXP	1	0	424	200	0	0	0	624	1.55	1.35
4070	EXP	1	0	189	96	0	0	0	285	16.09	15.62

MONTH = 7705 , FLAG = 3

1033	CNT	1	0	504	240	0	0	0	744	0.00	0.00
4008	EXP	1	0	207	72	0	0	0	279	3.65	4.70
4009	EXP	1	0	181	72	0	0	0	253	3.01	3.67
4019	EXP	1	0	391	216	0	0	0	607	3.64	3.11
4028	EXP	1	0	374	168	0	0	0	542	8.48	7.49
4039	EXP	1	0	230	72	0	0	0	302	3.65	4.32
4043	EXP	1	0	446	164	0	0	0	610	5.04	5.24
4051	EXP	1	0	445	216	0	0	0	661	3.85	3.38
4081	EXP	1	0	209	72	0	0	0	281	4.18	4.82

MONTH = 7707 , FLAG = 3

1084	CNT	1	0	480	264	0	0	0	744	0.00	0.00
4007	EXP	1	0	360	192	0	0	0	552	3.74	5.15
4020	EXP	1	0	480	264	0	0	0	744	0.00	0.00
4045	EXP	1	0	456	236	0	0	0	692	3.17	5.08
4063	EXP	1	0	323	144	0	0	0	467	4.00	1.98

MONTH = 7708 , FLAG = 3

1081	CNT	1	0	374	144	0	0	0	518	7.49	8.71
4007	EXP	1	0	156	48	0	0	0	204	5.06	4.07
4015	EXP	1	0	170	96	0	0	0	266	8.98	5.94

MONTH = 7709 , FLAG = 3

CUSTOMER ID	TREATMENT	STRATA	NUMBER OF ZERO KWH VALUES FOR RECORD TYPE							MAXIMUM HOURLY USAGE IN MONTH	
			1	2	3	4	5	6	7	WEEKDAY	WEEKEND
			---	---	---	---	---	---	---	-----	-----
1081	CNT	1	0	504	216	0	0	0	720	0.00	0.00
4003	EXP	1	0	413	144	0	0	0	557	3.64	5.99
4023	EXP	1	0	360	97	0	0	0	457	2.02	1.49
4035	EXP	1	0	237	48	0	0	0	285	6.26	9.18
4044	EXP	1	0	489	216	0	0	0	705	3.10	0.00
4061	EXP	1	0	504	216	0	0	0	720	0.00	0.00

* Record Type Definitions

- 1...Day of monthly system peak
- 2...Weekdays, excluding holidays
- 3...Weekend days, including holidays
- 4...All days where system load exceeds 0.80 times the assumed annual system peak
- 5...All days where system load exceeds 0.90 times the assumed annual system peak
- 6...All days where system load exceeds 0.95 times the assumed annual system peak
- 7...All days.

APPENDIX D: COMPARISON OF EXPERIMENTAL RESPONSES

UNDER TOU RATE AND CONTROL RATE

This appendix contains eight tables. Seven of these tables provide sets of summary statistics for the months of August, September, and December 1976 and January, July-September 1977. The 20 parameters identified in the first column of each table are described in Table D-1. A detailed description of each parameter was given in Section 3.2 of Chapter 3. Columns 2-4 of each table gives respectively, the sample size, parameter estimate and standard error of the estimate for the control group. Columns 5-7 give corresponding values for the experimental group. For both groups, sample results have been weighted to produce estimates and standard errors applicable to the effective population. Column 8 shows the difference in parameter estimates for the control and experimental groups. The standard error of this estimated difference is given in column 9. The hypothesis that the difference between the actual population parameters for the two groups is zero is tested and the result indicated by the standard notation of a single asterisk for statistical significance at the .05 level of significance or a double asterisk for significance at the .01 level. An approximate ninety-nine percent confidence interval for the true difference is given in the last two columns. This interval is constructed such that the chances are 99 times out of 100 that the interval will cover the true difference between the population parameters.

Table D-1. List of Parameters Estimated in Comparative Analysis
(Estimated per customer)

	Parameter Number	Description
Average for Day of Monthly System Peak	1	kWh - peak
	2	kWh - intermediate
	3	kWh - base
	4	kWh - during 1-hr. preceding peak period
	5	kWh - during 1-hr. following peak period
	6	Total kWh
Average for All Days of Month	7	kWh - peak
	8	kWh - intermediate
	9	kWh - base
	10	kWh - during 1-hr. preceding peak period
	11	kWh - during 1-hr. following peak period
	12	Total kWh
Average per Customer	13	Diversified demand at time of observed monthly system peak
	14	Noncoincident maximum demand
Average for Subset of Hours during Month	15	Average hourly kWh consumed during all hrs. of observed system load greater than 80% of <u>annual</u> system peak
	16	Average hourly kWh consumed during all hrs. of observed system load greater than 90% of <u>annual</u> system peak
	17	Average hourly kWh consumed during all hrs. of observed system load greater than 95% of <u>annual</u> system peak
Ratio	18	Proportion of total kWh used during peak period on peak day [#1/#6]
	19	Proportion of total kWh used during peak period on average day [#7/#12]
	20	Class load factor based on diversified demand at time of observed monthly system peak (#12/#13 ÷ 24)

Table D-2. Comparison of Experimental and Control Groups, August 1976

PARAMETER	PARAMETER ESTIMATES FOR CNT			PARAMETER ESTIMATES FOR EXP			TEST FOR TREATMENT DIFFERENCES			
	N	ESTIMATE	STD.ERR.	N	ESTIMATE	STD.ERR.	DIFFERENCE	STD.ERR.	99% CONF. LIMITS	
1	50	11.854	1.125	87	7.146	0.578	-4.709**	1.264	-7.966	-1.452
2	50	0.0	0.0	87	0.0	0.0	0.0	0.0		
3	50	24.430	1.584	87	27.512	1.648	3.082	2.286	-2.806	8.970
4	50	1.693	0.185	87	1.864	0.192	0.172	0.267	-0.515	0.858
5	50	2.117	0.237	87	2.231	0.180	0.114	0.298	-0.652	0.881
6	50	36.284	2.555	87	34.657	1.906	-1.627	3.188	-9.839	6.585
7	50	6.894	0.510	87	4.274	0.307	-2.620**	0.595	-4.152	-1.087
8	0			0						
9	50	24.024	1.431	87	25.451	1.391	1.427	1.996	-3.713	6.567
10	50	0.987	0.075	87	1.066	0.073	0.079	0.105	-0.190	0.349
11	50	1.325	0.112	87	1.339	0.095	0.013	0.147	-0.364	0.391
12	50	30.917	1.893	87	29.725	1.589	-1.192	2.471	-7.558	5.173
13	50	2.080	0.213	87	1.243	0.114	-0.837**	0.242	-1.460	-0.214
14	50	5.999	0.295	87	6.329	0.280	0.330	0.407	-0.718	1.378
15	50	1.838	0.143	87	1.479	0.098	-0.360*	0.173	-0.806	0.087
16	50	1.934	0.171	87	1.268	0.113	-0.665**	0.205	-1.192	-0.138
17	0			0						
18	50	0.327	0.013	87	0.206	0.014	-0.121**	0.019	-0.170	-0.071
19	50	0.223	0.006	87	0.144	0.007	-0.079**	0.010	-0.104	-0.054
20	50	0.619	0.054	87	0.997	0.092	0.377**	0.106	0.103	0.651

Table D-3. Comparison of Experimental and Control Groups, September 1976

PARAMETER	PARAMETER ESTIMATES FOR CNT			PARAMETER ESTIMATES FOR EXP			TEST FOR TREATMENT DIFFERENCES			
	N	ESTIMATE	STD.ERR.	N	ESTIMATE	STD.ERR.	DIFFERENCE	STD.ERR.	99% CONF. LIMITS	
1	50	8.066	0.640	88	5.965	0.672	-2.101*	0.928	-4.490	0.289
2	50	0.0	0.0	88	0.0	0.0	0.0	0.0		
3	50	20.970	1.303	88	22.754	1.465	1.784	1.961	-3.266	6.835
4	50	1.034	0.129	88	1.095	0.121	0.061	0.176	-0.393	0.516
5	50	1.809	0.190	88	2.113	0.207	0.304	0.281	-0.420	1.028
6	50	29.036	1.809	88	28.719	1.876	-0.317	2.606	-7.030	6.397
7	50	5.270	0.311	88	3.519	0.198	-1.751**	0.369	-2.700	-0.801
8	0			0						
9	50	21.423	1.259	88	21.983	1.088	0.560	1.664	-3.726	4.845
10	50	0.768	0.053	88	0.828	0.058	0.060	0.079	-0.142	0.262
11	50	1.079	0.068	88	1.105	0.068	0.026	0.096	-0.222	0.274
12	50	26.693	1.525	88	25.502	1.222	-1.191	1.954	-6.224	3.842
13	50	1.539	0.174	88	1.087	0.141	-0.452*	0.224	-1.028	0.125
14	50	5.631	0.270	88	5.833	0.254	0.202	0.371	-0.754	1.157
15	50	1.463	0.122	88	1.135	0.097	-0.328*	0.156	-0.730	0.074
16	0			0						
17	0			0						
18	50	0.278	0.011	88	0.208	0.016	-0.070**	0.020	-0.122	-0.018
19	50	0.197	0.006	88	0.138	0.005	-0.059**	0.008	-0.080	-0.039
20	50	0.723	0.078	88	0.977	0.110	0.255	0.135	-0.094	0.603

Table D-4. Comparison of Experimental and Control Groups, December 1976

PARAMETER		PARAMETER ESTIMATES FOR CNT			PARAMETER ESTIMATES FOR EXP			TEST FOR TREATMENT DIFFERENCES					
	:	N	ESTIMATE	STD.ERR.	:	N	ESTIMATE	STD.ERR.	:	DIFFERENCE	STD.ERR.	99% CONF. LIMITS	
1	:	0			:	0			:				
2	:	0			:	0			:				
3	:	0			:	0			:				
4	:	0			:	0			:				
5	:	0			:	0			:				
6	:	0			:	0			:				
7	:	49	8.847	1.086	:	92	6.977	0.630	:	-1.870	1.256	-5.104	1.365
8	:	0			:	0			:				
9	:	49	33.096	4.319	:	92	37.865	4.054	:	4.769	5.923	-10.488	20.026
10	:	49	1.338	0.189	:	92	1.426	0.146	:	0.088	0.239	-0.528	0.704
11	:	49	1.650	0.168	:	92	2.017	0.210	:	0.366	0.269	-0.327	1.060
12	:	49	41.943	5.388	:	92	44.842	4.494	:	2.900	7.016	-15.173	20.973
13	:	0			:	0			:				
14	:	49	6.814	0.541	:	92	7.022	0.469	:	0.208	0.716	-1.636	2.051
15	:	0			:	0			:				
16	:	0			:	0			:				
17	:	0			:	0			:				
18	:	0			:	0			:				
19	:	49	0.211	0.004	:	92	0.156	0.011	:	-0.055**	0.012	-0.086	-0.025
20	:	0			:	0			:				

Table D-5. Comparison of Experimental and Control Groups, January 1977

PARAMETER	PARAMETER ESTIMATES FOR CNF			:	PARAMETER ESTIMATES FOR EXP			:	TEST FOR TREATMENT DIFFERENCES		
	N	ESTIMATE	STD.ERR.		N	ESTIMATE	STD.ERR.		DIFFERENCE	STD.ERR.	99% CONF. LIMITS
1	0			:	0			:			
2	0			:	0			:			
3	0			:	0			:			
4	0			:	0			:			
5	0			:	0			:			
6	0			:	0			:			
7	51	8.556	1.024	:	90	6.739	0.690	:	-1.817	1.235	-4.998 1.364
8	0			:	0			:			
9	51	38.902	5.480	:	90	43.453	5.327	:	4.550	7.642	-15.135 24.236
10	51	1.337	0.175	:	90	1.446	0.167	:	0.108	0.242	-0.515 0.732
11	51	1.617	0.186	:	90	2.030	0.278	:	0.412	0.335	-0.450 1.274
12	51	47.458	6.484	:	90	50.192	5.792	:	2.734	8.694	-19.660 25.127
13	0			:	0			:			
14	51	6.734	0.521	:	90	7.365	0.560	:	0.631	0.765	-1.340 2.602
15	0			:	0			:			
16	0			:	0			:			
17	0			:	0			:			
18	0			:	0			:			
19	51	0.180	0.005	:	90	0.134	0.011	:	-0.046**	0.012	-0.078 -0.014
20	0			:	0			:			

Table D-6. Comparison of Experimental and Control Groups, July 1977

PARAMETER ESTIMATES FOR CNT				PARAMETER ESTIMATES FOR EXP				TEST FOR TREATMENT DIFFERENCES			
PARAMETER	N	ESTIMATE	STD.ERR.	PARAMETER	N	ESTIMATE	STD.ERR.	DIFFERENCE	STD.ERR.	99% CONF. LIMITS	
1	0			1	0						
2	0			2	0						
3	0			3	0						
4	0			4	0						
5	0			5	0						
6	0			6	0						
7	49	8.338	0.654	7	91	6.013	0.413	-2.325**	0.773	-4.316	-0.333
8	0			8	0						
9	49	32.651	2.133	9	91	32.938	1.779	0.286	2.778	-6.868	7.441
10	49	1.229	0.094	10	91	1.259	0.077	0.031	0.122	-0.282	0.344
11	49	1.502	0.119	11	91	1.430	0.086	-0.072	0.147	-0.452	0.307
12	49	40.989	2.739	12	91	38.951	2.049	-2.038	3.420	-10.848	6.772
13	0			13	0						
14	49	6.508	0.291	14	91	6.585	0.282	0.077	0.405	-0.967	1.121
15	0			15	0						
16	0			16	0						
17	0			17	0						
18	0			18	0						
19	49	0.203	0.005	19	91	0.154	0.007	-0.049**	0.009	-0.073	-0.025
20	0			20	0						

Table D-7. Comparison of Experimental and Control Groups, August 1977

PARAMETER	PARAMETER ESTIMATES FOR CNI			PARAMETER ESTIMATES FOR EXP	TEST FOR TREATMENT DIFFERENCES			
	N	ESTIMATE	STD.ERR.		DIFFERENCE	STD.ERR.	99% CONF. LIMITS	
1	0			0				
2	0			0				
3	0			0				
4	0			0				
5	0			0				
6	0			0				
7	49	7.648	0.563	93	5.840	0.365	-1.808**	0.671 -3.535 -0.080
8	0			0				
9	49	26.408	1.715	93	28.228	1.458	1.819	2.251 -3.978 7.617
10	49	1.174	0.088	93	1.197	0.077	0.023	0.117 -0.279 0.325
11	49	1.381	0.113	93	1.398	0.083	0.017	0.140 -0.344 0.378
12	49	34.056	2.243	93	34.068	1.711	0.011	2.821 -7.254 7.277
13	0			0				
14	49	6.259	0.282	93	6.406	0.288	0.147	0.403 -0.891 1.185
15	0			0				
16	0			0				
17	0			0				
18	0			0				
19	49	0.225	0.005	93	0.171	0.007	-0.053**	0.009 -0.076 -0.031
20	0			0				

D-8

Table D-7. Comparison of Experimental and Control Groups, September 1977

PARAMETER	PARAMETER ESTIMATES FOR CNT			:	PARAMETER ESTIMATES FOR EXP			:	TEST FOR TREATMENT DIFFERENCES		
	N	ESTIMATE	STD.ERR.		N	ESTIMATE	STD.ERR.		DIFFERENCE	STD.ERR.	99% CONF. LIMITS
1	0			:	0			:			
2	0			:	0			:			
3	0			:	0			:			
4	0			:	0			:			
5	0			:	0			:			
6	0			:	0			:			
7	49	6.269	0.470	:	90	4.536	0.255	:	-1.733**	0.535	-3.111 -0.355
8	0			:	0			:			
9	49	24.353	1.611	:	90	24.653	1.292	:	0.300	2.065	-5.020 5.620
10	49	0.954	0.082	:	90	0.931	0.064	:	-0.023	0.104	-0.292 0.245
11	49	1.160	0.086	:	90	1.171	0.067	:	0.012	0.109	-0.269 0.292
12	49	30.622	2.037	:	90	29.189	1.478	:	-1.433	2.516	-7.915 5.049
13	0			:	0			:			
14	49	5.950	0.275	:	90	6.104	0.257	:	0.154	0.377	-0.816 1.124
15	0			:	0			:			
16	0			:	0			:			
17	0			:	0			:			
18	0			:	0			:			
19	49	0.205	0.006	:	90	0.155	0.006	:	-0.049**	0.008	-0.070 -0.028
20	0			:	0			:			

D-9

Research Triangle Institute
P.O. Box 12194
Research Triangle Park, N.C. 27709
Telephone 919-541-6000