FACTORS AFFECTING RESIDENTIAL HEATING ENERGY CONSUMPTION

John J. Donovan
Walter P. Fischer

Sloan School Center for Information Systems Research
in association with the Energy Laboratory

Working Paper No. MIT-EL-76-004WP

FACTORS AFFECTING RESIDENTIAL HEATING ENERGY CONSUMPTION[1]

John J. Donovan[2]
Walter P. Fischer[3]

April 1976
this revision July 16, 1976

CONTENTS

---

[1] This work was supported in part by the New England Regional Commission as part of the NEEMIS Project, Contract No. 10630776 and by the M.I.T./IBM Joint Study Agreement. The Scott Oil Company greatly assisted in the data gathering process.

[2] Massachusetts Institute of Technology

[3] University of Munich and an IBM Research Fellow

ABSTRACT

Twenty per cent of all energy consumed in New England and ten per cent of all energy consumed in the United States is consumed in home heating. This paper reports on an effort to ascertain the major factors affecting the consumption of home heating oil. Three general classes of factors are analyzed: (1) physical and occupant characteristics (number of rooms, number of occupants, number of stories, amount of insulation, income level, etc.); (2) external (price, shortage awareness, weather); and (3) behavioral and physical changes (change in temperature settings, change in insulation, change in oil burner, etc.).

The study is based on four data series: (1) actual monthly home heating oil consumption data on 8000 suburban homeowners in suburban Boston; (2) questionnaire responses from 2000 homeowners on their homes' physical and occupant characteristics, as well as changes in physical and occupant behavioral characteristics between 1972 and 1975; (3) monthly weather data; and (4) heating oil price data. The data is associated with the years from 1972 through 1975, a period in which marked price changes, shortages, and behavioral changes occurred, hence providing an opportunity to study the effects of these various events.

Three models are central to the study:

Model I. A cross-sectional model that depicts consumption per degree-day as a function of physical and occupant characteristics of a home.

Model II. A time series regression model that establishes consumption per degree-day as a function of price and

consumer awareness of an energy shortage.

Model III. A cross-sectional regression model that attempts to explain change in consumption per degree-day from one year to the next as a function of specific conservation actions such as temperature resetting, addition of storm windows, etc.

The major findings of each model are as follows:

Model I: House size, age of home, family income, and the presence of storm doors and windows are all significant factors in predicting the amount of home oil consumption.

Model II: Estimated values of price elasticity with respect to demand for residential heating oil and a measure of impact of shortage awareness on consumption are determined. This model also demonstrates that there were substantial savings in consumption corresponding to increases in price and shortage awareness from 1972-1975.

Model III: The data from the questionnaire indicate that only a few consumers made physical home improvements; however, the data from the oil company indicate that a substantial savings (over 12%) in consumption occurred between the heating seasons 1972/73 and 1973/74. The conclusion from this data indicates that behavioral changes were the major conservation actions taken. Model III indicates that the behavioral change of temperature resetting is significant and the physical change of additional weather stripping and change of burner are significant. Further study is needed, however, to determine those behavioral changes that accounted for the major change in consumption. In addition, this model indicates that different groups within the sample (e.g., by income level, house characteristics) display similar conservation efficiency.

In addition to the findings of the models, the paper includes (in Appendix B) a detailed discussion of biases associated with the data. Major conclusions from that discussion are: (1) our sample is representative of suburban homes in the Northern United States; (2) the consumers who responded to the questionnaire were slightly more energy-conscious and responded slightly more dramatically to price increases than the general populace; (3) our residential heating oil prices are representative of those that prevailed in the region; and (4) the heating seasons 1972 through 1975 were warmer than usual. Trends in the data indicate that new homes in the sample have a considerable amount of insulation and the typical single-family house in the sample has storm windows and doors.

1. INTRODUCTION

Residential space heating consumes over 20 percent of all energy used in New England[4] and comprises over 10 per cent of all energy consumed in the United States.[5] Oil is the source of over 70 percent of New England's home heat and virtually all of this oil is imported into the region.[6] Hence, even a small reduction in home heating oil consumption could result in considerable savings to the region. The New England Regional Commission has estimated that a 5 percent reduction of energy consumed by homes in New England would result in a net savings of $87.5 million to the region. This study indicates such a decrease in consumption is seemingly attainable. However, efforts to determine, encourage, quantify, and sustain behavioral conservation actions must be made, while at the same time promoting physical home improvements.

This study attempts to determine and quantify those factors affecting home heating oil consumption. The report focuses on single-family, suburban, homes where weather patterns are similar to those in all of New England. Single-family suburban homes are a group that warrants study, as over 50 percent of all housing units in New England and 60 percent of all in the United States are single-family dwellings, and over 50 percent of all Americans live in a suburb.[7]

---

4 Preliminary Projections of New England's Energy Requirements, prepared for New England Regional Commission (NERCOM) by Arthur D. Little, Inc., 1975.

5 Stephen Dole, "Energy Use and Conservation in the Residential Sector: A Regional Analysis," Rand Corporation report, R-1641-NSF.

6 "Fuel Trade and Fact Book," Yankee Oil Man, March 1974.

7 Detailed Housing Characteristics, U.S. Department of Commerce, Social and Economic Statistics Administration, Bureau of the Census, 1970, HC(1)-B1 and HC(1)-B23.

2. DATA: SOURCES AND CONTEXT

The major data series used in this study come from four sources:

(1) Consumption data on individual homes--oil delivery records for some 8000 single-family homes in suburban Boston for the heating seasons (December through March) 1972/73, 1973/74, and 1974/75.[8]

(2) Questionnaire data--responses from 2000 customers to a questionnaire sent in December 1974 (see Appendix A).[9]

(3) Weather data--monthly weather data for 38 locations throughout New England. Seven of these locations had detailed monthly weather histories of nearly forty years.[10]

(4) Price data--monthly price of oil to the end-user.[11]

---

[8] We acknowledge those at Scott Oil Company, Boston, and especially Mr. Thomas J. Scott, who personally made available consumption data and through the facilities of his corporation sponsored and carried out the entire effort of collecting data on house characteristics by means of a questionnaire; Mr. Harvey Deitel of Business Computer Services, Needham Heights, Massachusetts, who extracted consumption data from Scott Oil Company records. We thank those consumers who voluntarily responded to the questionnaire.

[9] 18,000 questionnaires were sent out to five distributors of Scott Oil. 2,000 responded. From one of these distributors we obtained oil delivery records for roughly 8000 customers. The useful intersection between these oil delivery records and the questionnaire responses was 668.

[10] Degree-day information was provided by Weather Services Corporation, Bedford, Massachusetts.

[11] This data was supplied by Scott Oil Company for the heating seasons 1972/73, 1973/74, and 1974/75, and other series obtained from "The Analysis of the Impacts on New England of Recent Energy Shortages and Price Increases" prepared by Ernst & Ernst for the New England Regional Commission, Boston, Massachusetts (January, 1975) and the National Energy Review.

The total data set is unusual; it provides both actual consumption data from the company delivering oil to a home and also data from the homeowner on physical characteristics of his home as well as the behavioral and personal characteristics of the occupants of that home. Compiling both consumption and characteristic data has provided a vehicle from which meaningful insights and inferences can be drawn. In general, other studies employ one or the other, but rarely both of these data series, as their sources of data.[12,13,14] Further, the data spans a period of dramatic changes in price, consumption, and consumer awareness of energy shortages, as it includes the 1973/74 heating season during which the Arab oil embargo and related events caused a highly volatile market. Hence, it offers a rare opportunity to analyze the change in consumption associated with these fluctuations in price and awareness.

---

12 Robert Perlman and Roland L. Warren: Energy-Saving by Households in Three Metropolitan Areas, Report #1 of the Energy Impact Study, Brandeis University, Waltham, Massachusetts, March 1975.

13 1976 National Energy Outlook, FEA/N-76/100, Stock No. 041-018-00097-6.

14 Peck & Doering: "Voluntarism and Price Response: Consumer Reaction to the Energy Shortage," Bell Journal, Spring 1976.

3. DATA: SAMPLING, PROCEDURES, AND SUMMARY STATISTICS

Before directing attention to the major focus of this study (using models to determine factors affecting residential heating oil consumption), we briefly discuss a few points pertaining to physical aspects of households and trends that were observed while examining the raw data.

Figure 1 puts into perspective the sample data used in this section.

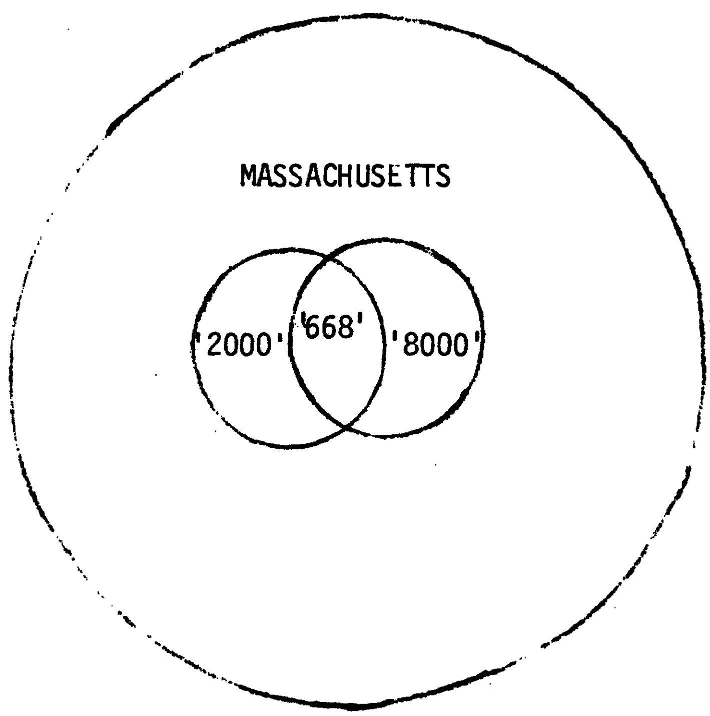


Figure 1: Subsample Used in Trend Analysis

The set Massachusetts denotes all homeowners in Massachusetts. The set '8000' denotes homeowners that receive oil from Scott distributors, and for whom detailed delivery data for the heating seasons studied was supplied. It can be assumed that for single-family houses all residential fuel delivered is consumed (in some cases oil is also used to heat hot water, however, this represents a very small percentage of fuel consumed during the heating season). Hence, consumption data can be calculated from delivery data. The set '2000' denotes consumers that responded to the questionnaire. The meaningful intersection of these two sets is the set '668', which is composed of consumers for which data on home and occupant characteristics as well as on consumption is available.

To discuss the trends observed in the new data, one must ask if the data is representative of the region. In Appendix B we show that the consumption rate of the subsample '668' is similar to that of the sample '8000' (differing by only 3%). Hence, it is possible to infer that this subsample has energy improvement characteristics similar to the sample '8000' which is representative of suburban homes in the region.

Figure 2a shows that newer homes in the sample tend to have more insulation and fewer stories. Figure 2b reveals that a relatively high percentage of homes in the sample are fully equipped with storm doors and windows. This observation is consistent with that of Newman and Day on a national basis.[15]

[15] Dorothy K. Newman and Dawn Day, The American Energy Consumer, a report to the Energy Policy Project of the Ford Foundation, Ballinger, Cambridge, Massachusetts, 1975.

Year House Was Built	% of Sample	% Insulated Walls	% Insulated Ceilings	Average # of Rooms	Average # of Stories
1901 to 1910	4	53	35	7.8	2.5
1910 to 1920	4	48	44	8.2	2.5
1921 to 1930	8	34	32	7.2	2.2
1931 to 1940	16	36	44	6.8	2.1
1941 to 1950	20	60	54	6.5	2.1
1951 to 1960	24	81	75	6.3	1.6
1961 to 1970	19	87	82	7.0	1.7
1971 to 1974	5	92	86	8.1	1.9

Figure 2a: Distribution by Age, Percentage of Houses Having Insulated Walls and Ceilings by Age Group, Average Number of Rooms, Average Number of Stories

Type of Equipment	Fraction of the House Equipped				
	0	1/4	1/2	3/4	1
Weather stripping	39	6	17	3	34
Storm windows	8	4	5	8	75
Storm doors	9	2	4	5	79

Figure 2b: Percentage of Houses with Weather Stripping, Storm Doors, or Storm Windows

4. MODELS

Three regression models were constructed to quantify and predict the effect of physical, behavioral, and external changes on consumption.[16] To normalize for the effects of weather, consumption of a home is expressed as gallons of oil consumed per degree-day (CPD). Degree-days are a weighted average of daily temperature as they vary from a mean of 65°. All three models use a variant of consumption per degree-day as their dependent variable:

Model I--a cross-sectional regression model that expresses average yearly consumption per degree-day for each consumer (CPD_y) as a function of physical and occupant characteristics:

$$CPD_y = f(\text{no. of rooms, age of home, income, no. of occupants, etc.}).$$

This model uses data from the 668 persons for whom both consumption information and physical/occupant characteristics are available.

Model II--a time series econometric regression model that relates average monthly consumption per degree-day averaged over all consumers ($\overline{CPD}_m$) as a function of external influences:

$$\overline{CPD}_m = f(\text{price, shortage awareness})$$

This model uses data from the 300 persons for whom monthly consumption data could be determined.

Model III--a cross-sectional, regression model that attempts to explain changes in the average yearly consumption per degree-day for each consumer

---

[16] We acknowledge the help of Bruce Stangle in the organization and several reviews of this work.

(ΔCPD_y) as a function of the low-level mechanisms that may have caused the change:

$$\Delta CPD_y = f(\text{change in thermostat settings, change in oil burner, change in insulation, etc.}).$$

This model uses data from the 668 persons for whom consumption, physical, occupant, and conservation information were available.

Both Model II and Model III relate changes in consumption to some factors that influence consumption. Model II relates consumption to economic factors (e.g., price and awareness), while Model III relates consumption to more engineering-oriented factors, such as temperature resettings and changes in burners. Model II will show that price increases and awareness correspond to reductions in consumption. The mechanisms employed by homeowners to achieve these lower consumption levels are studied in Model III, which relates the observed changes in consumption to the mechanisms, in an effort to give insights into which mechanisms should be encouraged on a broad scope.

Analysis of this kind necessarily leads to concerns about biases in the data, multicolinearities, and unrecorded facts. Biases are summarized as data is presented in the discussion of each model. A complete discussion of all bias analysis is presented in Appendix B. Multicolinearities have been addressed by testing many different combinations of variables for each model, looking for stable relationships. Generally recognized relevant factors affecting consumption[17,18,19,20] are the underpinnings of the data studied. These factors were used in hopes of minimizing the number of unrecorded data items that would have proven significant.

An advanced computational facility made this study possible,[21] in that it facilitated the calculations of the different variants of CPD, validation of the data, analysis of biases (Appendix B), and the building of the models. This facility is discussed in detail elsewhere.[22,23]

---

17 R. Schoen, A. Hirshberg, and J. Weingart, New Energy Technologies for Buildings, a report to the Energy Policy Project of the Ford Foundation, Cambridge, Mass., Ballinger, 1975.

18 G. Dallaire, "Designing Energy-Conserving Buildings," Civil Engineering, vol. 44, no. 4, 54-58, New York, April, 1974.

19 Design and Evaluation Criteria for Energy Conservation in New Buildings, U.S. Department of Commerce, National Bureau of Standards, NBS Technical Note 789, July, 1973.

20 P. K. Schoenberger, "Energy Saving Techniques for Existing Buildings," Heat., Piping Air Conditions, Ohio State University, Columbus, vol. 47, no. 1, 98-105, January, 1975.

21 Recognition is given to an IBM/M.I.T. Joint Study that made available the computational facility of the IBM Cambridge Scientific Center. The software used in this work was jointly developed by IBM and M.I.T.

22 J. J. Donovan and H. D. Jacoby, "An Experimental System for Data Management and Analysis," M.I.T. Sloan School Center for Information Systems Research Report CISR-15, 1975.

23 J. J. Donovan, "DataBase Approach to Decision Support," ACM Transactions on Data Base Systems, vol. 3, December 1976.

5. CROSS-SECTIONAL MODEL OF CONSUMPTION DIFFERENCES BETWEEN HOUSES

One objective of this study was to establish a functional relationship between home/occupant characteristics and residential consumption of heating oil. The effects of physical characteristics (e.g., insulation, size, storm windows and doors, etc.) on a home are determined by engineering studies through the use of laboratory homes and/or appropriate computer simulations. Hence, they may not reflect the actual consumption patterns of homes lived in by representative consumers. An alternative is to use actual consumption data and data on characteristics of homes and occupants, as done in this study.

The approach used in this section is formulated around a cross-sectional regression model that draws from three data series: individual home consumption of heating oil, questionnaire responses, and weather data.

Figure 3 shows distributions of CPD per household in 1972/73, 1973/74, and 1974/75 for each home responding with an answered questionnaire. From this figure one can see that CPD for homes in the sample vary over a wide range (e.g., for 1973/74 the range was from .08 to .36). Model I attempts to determine the characteristics of homes and occupants that account for such a range.

The dependent variable (CPD_y) is the CPD per home during one heating season (where y = 1972/73, 1973/74, or 1974/75) and is formulated as a linear combination of household characteristics:

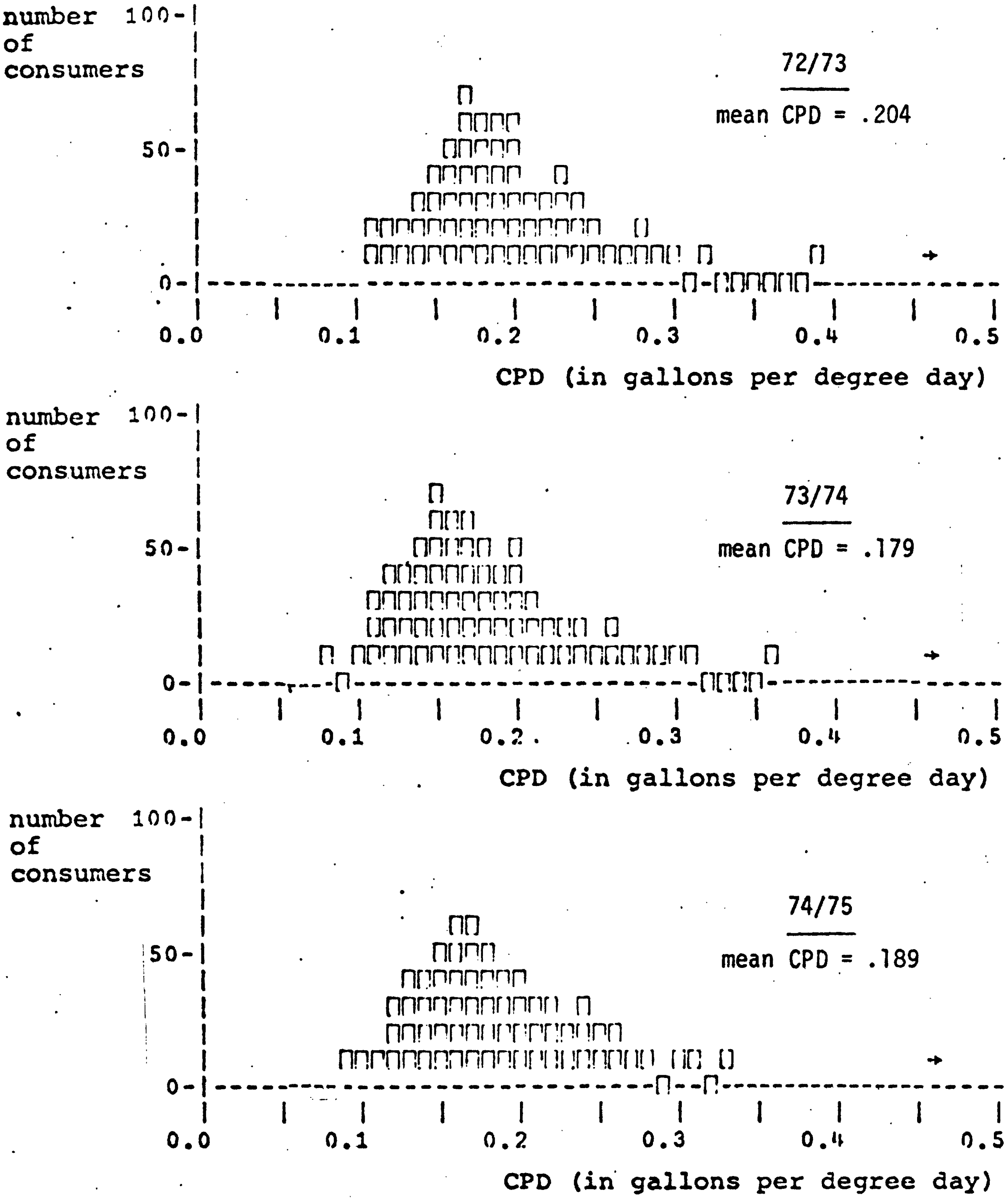


Figure 3: CPD Distributions for the Heating Seasons 1972/73, 1973/74, and 1974/75.

Model I: $$CPD_y = A_1 + \sum_{i=2}^{n} A_i X_i$$

where the X_i, i = 2, 3, ..., n are variables that characterize consumers and the A_i, i = 1, 2, ..., n are coefficients that are estimated by the ordinary least squares method.

Average $\overline{CPD}_y$ is calculated using delivery data as follows: (1) calculate consumption per degree-day for each day for a consumer (CPD_D) by dividing the amount of fuel delivered in a delivery period by the total degree-days in that period (a delivery period of the interval of time between deliveries). Hence, a CPD_D for all days in the winter months is calculated; (2) Consumption per degree-day for a year for each customer (CPD_y) is calculated by summing all the CPD_D for the period of December through March.

Note in all three models consumption has been normalized by degree days. This assumes that consumption is a linear function of degree-days. There is some engineering evidence that consumption has a quadratic or higher order relationship with weather. An alternative model would have weather as a right-hand side variable. However, from the analysis of bias we have found that in the range of degree-days considered in the winters under investigation, that consumption is such a linear function.[24]

[24] We acknowledge Professor Jerry Hausman for his suggestion to investigate this property of CPD.

The number of observations used in any given formulation of the model is dependent upon the variables involved since incomplete or missing observations were eliminated. In all cases the observations are a subset of the 668 persons for whom both consumption and physical data exists. Conclusions from the bias analysis reveal that CPD_y is slightly lower for the '668' subsample than for the populace as a whole (Section B.1, hypothesis #1) and that subsamples of '668' do not introduce any additional bias (Section B.1, hypothesis #3).

Figure 4 depicts the resulting coefficients (A_i), the associated standard errors and t-statistic for the constant term, and sixteen variables used to predict CPD in 1974/75.

Data used for these variables comes from the characteristics provided by the questionnaire for the '668' subsample. A conclusion of the bias analysis is that home characteristics of this '668' subsample typify characteristics of suburban homes. The following is a discussion of the independent variables used in the model. All variables were chosen to represent engineering or occupant characteristics that have been found in the literature to influence consumption.[25]

The variables X_2 and X_3 were chosen as indicators of size with X_2 being a linear term and X_3 being a quadratic term. The most accurate size indicator is the volume (e.g., in cubic feet) of the home. However, if a procedure for calculating the house volume had been part of the questionnaire, one would have expected (a) a lower response rate and (b) additional bias with regard to the type of person who

---

[25] See footnotes 17, 18, 19, and 20.

A_i	Term (X_i)	Coefficient	Standard Error	T-statistic
A_1	constant	.45137	.12871	3.50699
A_2	no. rooms + no. stories - 1	-.02465	.00644	-3.82938
A_3	(no. rooms + no. stories - 1)^2	.00229	.00035	6.54522
A_4	year home built	-.00039	.00010	-3.87211
A_5	storm windows	-.00625	.00233	-2.68756
A_6	storm doors	-.00318	.00231	-1.37781
A_7	income ($25,000-50,000)	.02484	.00676	3.67620
A_8	income (>$50,000)	.07837	.01245	6.29603
A_9	presence of basement	-.01430	.01360	1.05132
A_{10}	presence of insulation	-.00081	.00062	-1.30246
A_{11}	day temperature	.00130	.00136	0.95625
A_{12}	night temperature	.00082	.00076	1.07826

Number of observations	347
F-statistic	43.03893
Standard error	0.04832
r^2-statistic	0.58562

Figure 4: Model I Consumption in 1974/75 as a Function of Home and Occupant Characteristics

answered -- and possibly (c) inaccurate data. The same argument holds for other more accurate variables such as square footage of window area and amount and type of insulation.

There is a trade-off between the number and sophistication of questions asked and the response rate and bias of the responses. In the present study we chose simply to ask for the number of rooms and used this as an indication of the home's size. To compensate for deficiencies that result from using the number of rooms as an indicator of house size, we have introduced a quadratic term. This quadratic relationship was felt to be representative, as the number of rooms per house seems to be positively correlated with room size, thus calling for a non-linear relation (i.e., as the number of rooms goes up, so does the size of each room-- as large homes tend to have more and larger rooms).

Since we explicitly asked that halls, bathrooms, and stairways not be counted as rooms, one room was added to the indicated number of rooms for each two-story home, two rooms for each three-story home, and three for each four-story home, in order to account for the stairways and halls.

The variable X_4 represents the year the home was built. The variable X_5, storm windows, is given values 0 to 4 according to how many quarters of the house were so equipped. Similarly, X_6, storm doors, is a variable with four values. The variables X_7 and X_8, family income, are treated as two binary variables--one for the annual income group of $25,000 to $50,000 and one for the group above $50,000. The variable X_9 is a binary variable indicating presence (1) or absence (0) of a basement. The variable X_{10} is a binary variable indicating presence (1) or absence (0) of insulation. Variables X_{11} and X_{12} denote the absolute temperature setting in degrees Farenheit of daytime and nighttime setting respectively.

In Figure 4 all variables have the sign expected. For example, those that denoted influences that would lower consumption had a negative sign. The independent variables that were found to be significant are: number of rooms and stories, year house was built, presence of storm windows and doors, and income level.

Figure 5 depicts our best estimate of the influence of number of rooms and number of stories (Variables X_2 and the quadratic term X_3). The shape of the curve confirms our expectations. For homes in which number of rooms is low, the ratio of number of rooms to outside walls is low, hence, higher heat loss. Further, in homes with very few rooms, the rooms tend to be larger. For example, a four-room home is more likely to have a combination kitchenette/sitting room/family room, where a six-room home may have six relatively small, separate rooms. Hence, the relative flat portion from 4 to 7 units on the horizontal axis of Figure 5. However, as the number of rooms becomes larger, the room size becomes larger. (For example, a twelve-room house tends to have a much larger living room than a six-room home.) Hence, consumption rises as volume increases. Adding to this rise is the fact that more energy-demanding life-styles are often associated with owners of large homes.

Figures 6 and 7 depict our "best" estimate of the influence of three differing variables (X_4, X_7, and X_8) upon CPD based on the model results in Figure 4. Figure 6 confirms the view that more modern homes are associated with lower CPD. Figure 7 supports the observation of Newman and Day that high income levels are associated with higher CPD (X_7 and X_8). An explanation of this is that the higher income groups have a more energy-demanding lifestyle -- or that certain unrecorded home characteristics (e.g., room size) are markedly different.

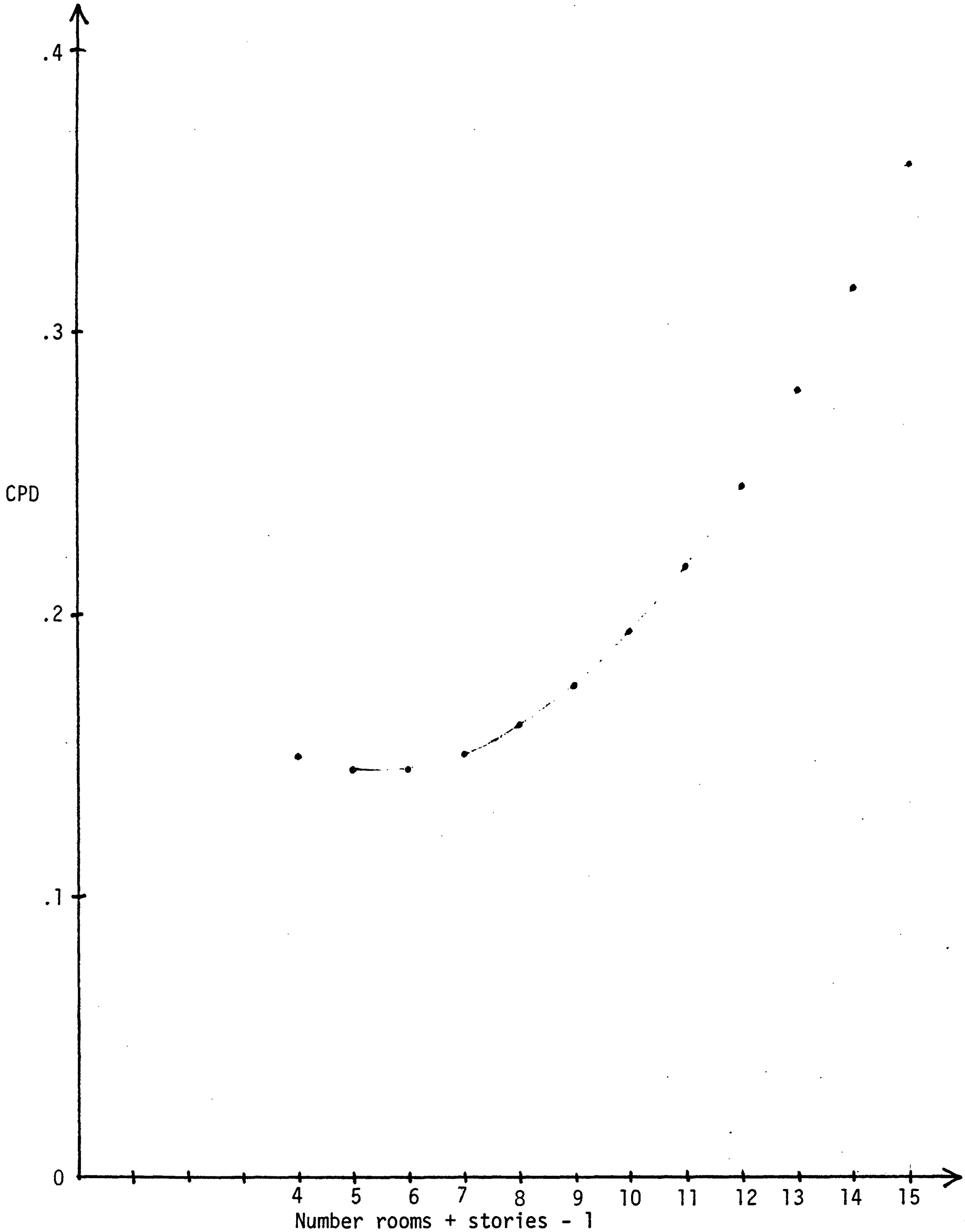


Figure 5: CPD for Single Family Earning Less than $25,000/year with a Basement. All other variables are set at their mean, while varying (rooms + stories - 1) from 4 to 15.

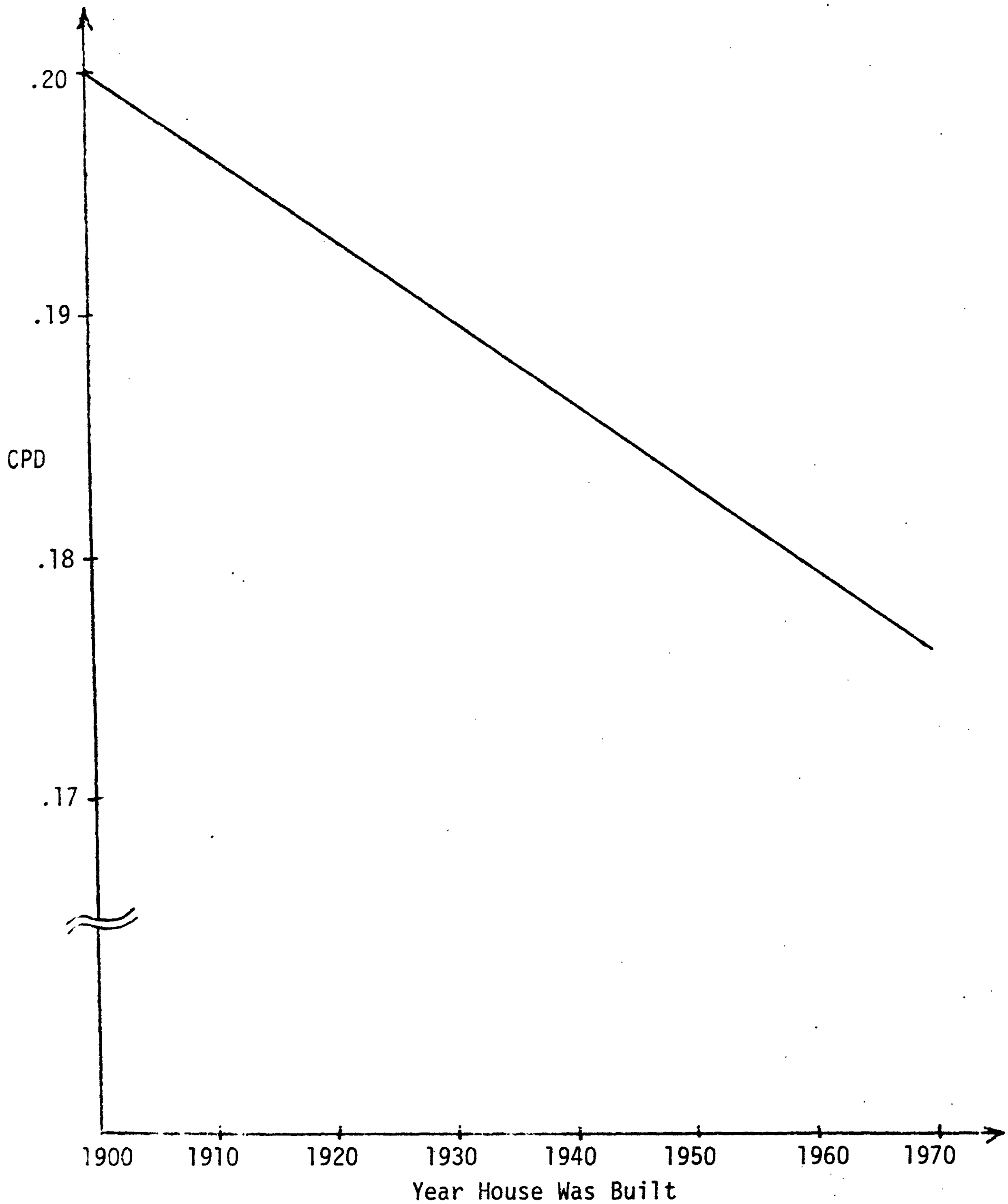


Figure 6: Consumption for Single Family Earning Less than $25,000 per Year and Living in Two-Story, Six-Room House and without Storm Doors or Windows

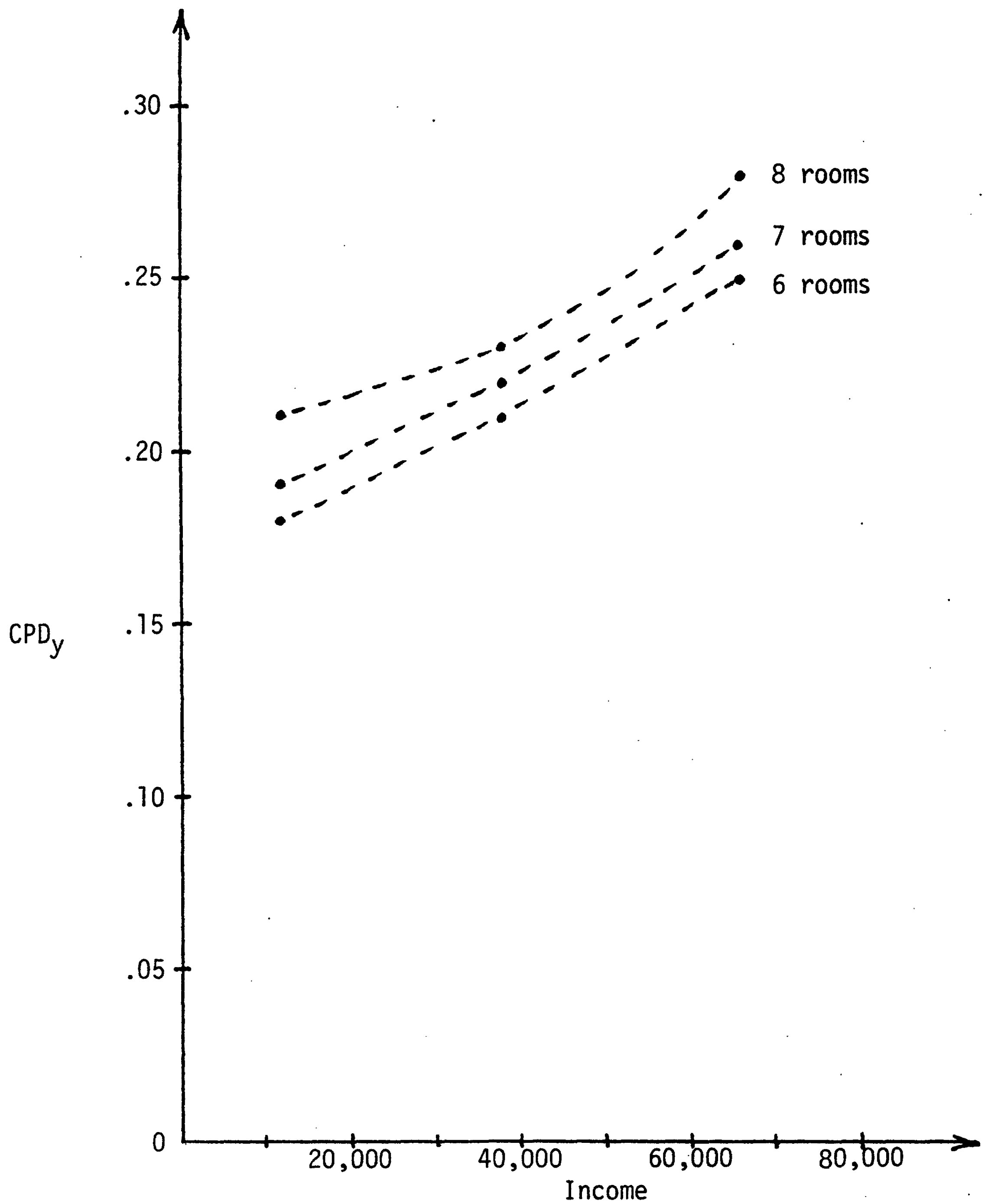


Figure 7: Consumption for Single Family Living in Two-Story House Built in 1953 and Having no Storm Windows or Doors

Using the model in Figure 4 and setting all variables to their mean values, except storm windows and storm doors X_5, X_6, the existence of storm doors corresponds to a 4.4% savings in consumption, and the existence of storm windows in the entire house corresponds to a 11% decrease in consumption over homes that have no storm windows.

In view of this study's limited ability to capture all relevant factors, an r^2 statistic of 0.58 is perceived as a satisfying result.

The variables (X_9, X_{10}, X_{11}, X_{12}) were recorded and subjected to analysis, and our best point estimate of their coefficient is given in Figure 4. Unfortunately, the data we have is not rich enough to give a high t-statistic. Variable X_9 corresponds to the existence of basements. While we feel, as do others (Newman and Day), that basements are important, our sample data did not have sufficient variations of this characteristic to show this statistically. Specifically, 96% of the sample have a basement or crawl space, 88% have storm doors (variable X_6), and 89% have storm windows (variable X_5), which may account for the low t-statistic of these variables.

The variable X_{10} denotes the presence of insulation, a factor confirmed by engineering studies to have an effect on consumption.[26] Apparently, however, either the knowledge of the insulation characteristics of an occupant's home was lacking or the phrasing of the question was confusing since one-fourth of the sample did not answer the question on the existence or nonexistence of insulation.

Engineering studies have shown absolute temperature setting (X_{11}, X_{12}) of the thermostat has an effect on consumption. However, the thermostat may be located in different locations in different homes (e.g.,

---

[26] John C. Moyers, The Value of Thermal Insulation in Residential Construction: Economics and the Conservation of Energy, Oak Ridge, National Laboratory, ORNL-NFS-EP-9, December, 1971.

some on outside walls, some on inside walls). Further, all thermostats in our sample were not calibrated in the same way. Hence, we are not surprised at the poor statistical information obtained form this variable (with this method of collecting its values).

Other recorded variables that were statistically insignificant and that do not have sufficient variations, and that may influence consumption (but with less certainty than those depicted in Figure 4), were the presence of an attic, attachment to another house, and type of material of which the house is constructed. In the sample 84% have an unheated attic, 83% are not attached to another house (99% were not attached on two sides to another house), and 83% are built of wood. Hence, these variables do not have much variation.

6. PRICE AND AWARENESS MODEL

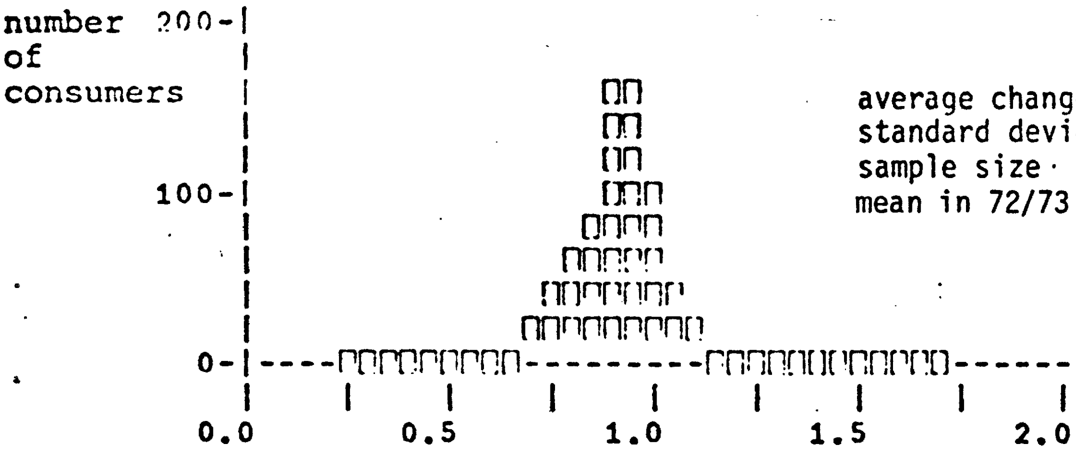
Another objective of this study was to determine the elasticity of demand with respect to price and awareness. The consumption data of the sample covers a period when there were perhaps the greatest price changes in recent history (1973/74 shows a 50% price increase) and the greatest awareness period of energy use, shortages, and expected price increases. Hence, the data affords an unusual opportunity to calculate short-term elasticities.

The consumption data from the sample demonstrates a marked change corresponding to changes in prices and public awareness of a fuel crisis (e.g., a savings of over 12% in CPD between the heating seasons of 1972/73 and 1973/74). This savings can be seen in Figure 8, where the horizontal axis indicates the fraction of 1972/73 CPD consumption in each of the following years. Note that if there had been no change in consumption, every entry would be equal to 1. Those consumers who used less oil are below 1, and those who consumed more are above 1. The first graph shows the saving tendencies during the energy crisis; the second shows how consumers returned to higher consumption in 1974/75; and the last shows that consumption in 1974/75 had not returned completely to the 1972/73 levels.

To analyze consumption as a function of factors that varied over time, a regression model was established:[27]

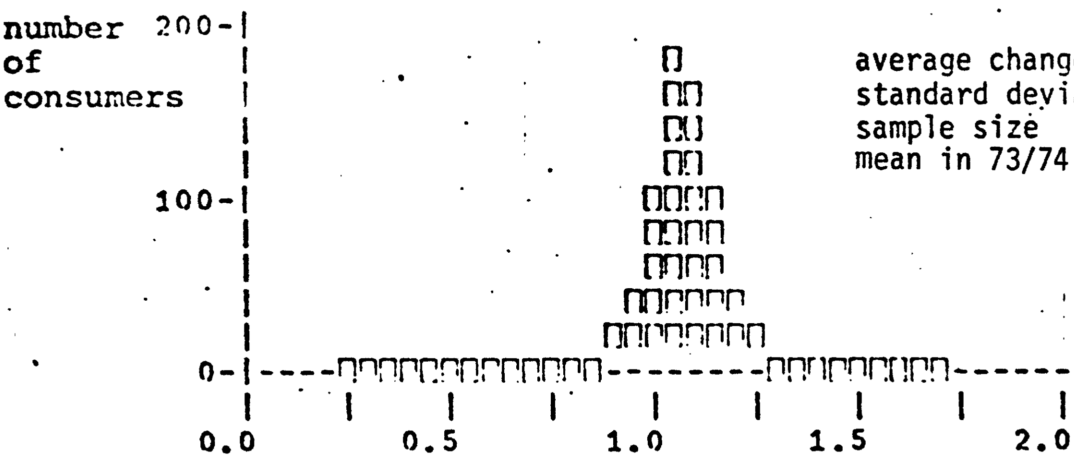
Model II: $$\overline{CPD}_m = A_1 + \sum_{i=2}^{n} A_i Z_i$$

[27] We acknowledge David O. Wood of the M.I.T. Energy Laboratory for assistance in the formulation of this model. We acknowledge the efforts of Richard Tabors and Henry Jacoby in assisting with the formulation, and Peter DiGiammarino for its implementation. We acknowledge Jerry Hausman's critical review.



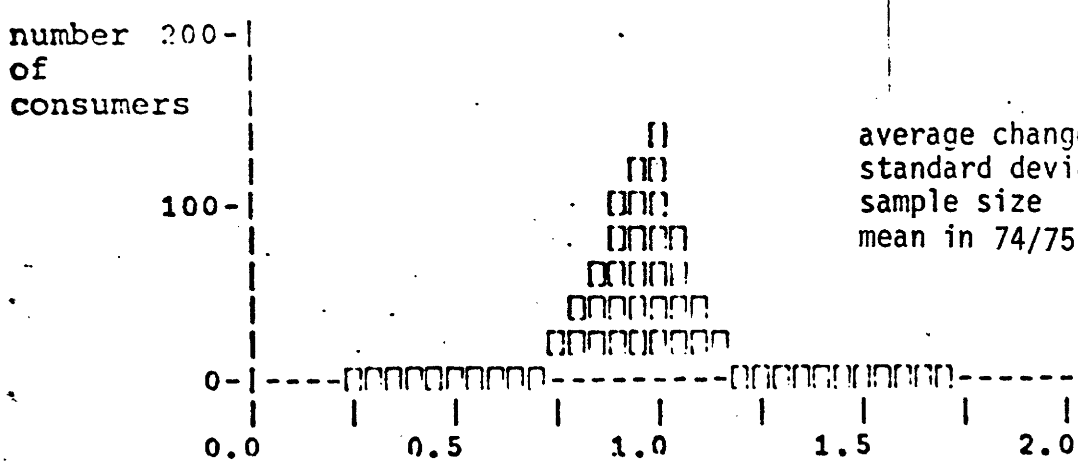
average change	0.882
standard deviation	0.101
sample size	668
mean in 72/73	.204

8a: **fraction of the 72/73 CPD consumed in 73/74**



average change	1.055
standard deviation	0.102
sample size	668
mean in 73/74	.179

8b: **fraction of the 73/74 CPD consumed in 74/75**



average change	0.927
standard deviation	0.117
sample size	668
mean in 74/75	.189

8c: **fraction of the 72/73 CPD consumed in 74/75**

Figure 8: Histograms of the Change in CPD between Heating Seasons

The dependent variable is the average CPD per month for all consumers of the '8000' sample who received frequent oil deliveries (five or more each season) during the three heating seasons (subset '300').

Data for the dependent variable ($\overline{CPD}_m$) was calculated using consumption records of the '300' consumers by the following procedure: (1) consumption for individual delivery periods was calculated for each customer using delivery data; (2) consumption per degree-day (CPD) for each customer was calculated by dividing the degree-days for each delivery period into the consumption of that period; (3) the average consumption per degree-day for all customers $\overline{CPD}$ for a particular day was obtained by averaging CPD for each customer for that day; and finally (4) the average consumption per degree-day for each day of a month $\overline{CPD}_m$ was calculated by summing $\overline{CPD}$ for each day of a month and dividing by the number of days.

A conclusion of the bias analysis is that subsample '300' reduced consumption by the same amount as did the general populace during the heating season 1973/74 but went back up by a lesser amount than the '8000' during the heating season 1974/75. Hence, one may expect a higher price elasticity and impact of awareness from this subsample than for the population as a whole.

The independent variables (X_2 and X_3) used in Model II are price and awareness:

X_2 Price was set equal to the average price of the oil company involved during the corresponding month (in cents/gallon). A discussion of bias in this data is presented in Appendix B, Section 3, with a conclusion that this price information is representative of New England. See Figure 9 for price changes used in this model.

X_3 Cumulative shortage awareness was set equal to a function of the number of front-page headline columns of energy-related articles in the Thursday and Sunday Boston Globe.[28] Values ranged from 0 to 60 per month. This was felt to be representative of the information available to homeowners. This variable was used to represent both the cumulative effect of such columns and the gradual decay of awareness with time after columns were published. A one-month lag between the presence of column and consumption changes was thought appropriate. Further, since the first column has the most substantial effect with each subsequent column having a lesser impact, the log of the total was used. Hence, the independent variable used to calculate the model results shown in Figure 10 is:

$$\text{CUMMAWARE}_I = \text{AWARE}[I-1] + .6 * (\text{CUMMAWARE}[I-1])$$

where I ranges over the months studied and .6 is the decaying effect, and AWARE [I] equals the number of front-page columns in the Thursday and Sunday Boston Globe in month I. Awareness is lagged two months. That is, the effect of awareness on consumption in month I is associated with the number of articles in month I-1. A plot of this function for the heating seasons 1972 through 1975 is depicted in Figure 10. Note values of awareness were obtained for all months in order to facilitate accurate calculations of cumulated awareness. For example, effect in December depends on awareness values in October.

---

28 We acknowledge Richard Tabors of Project NEEMIS and the M.I.T. Energy Laboratory for suggesting this measure of awareness. We acknowledge the suggestions of Jerry Hausman for making this a cumulative variable.

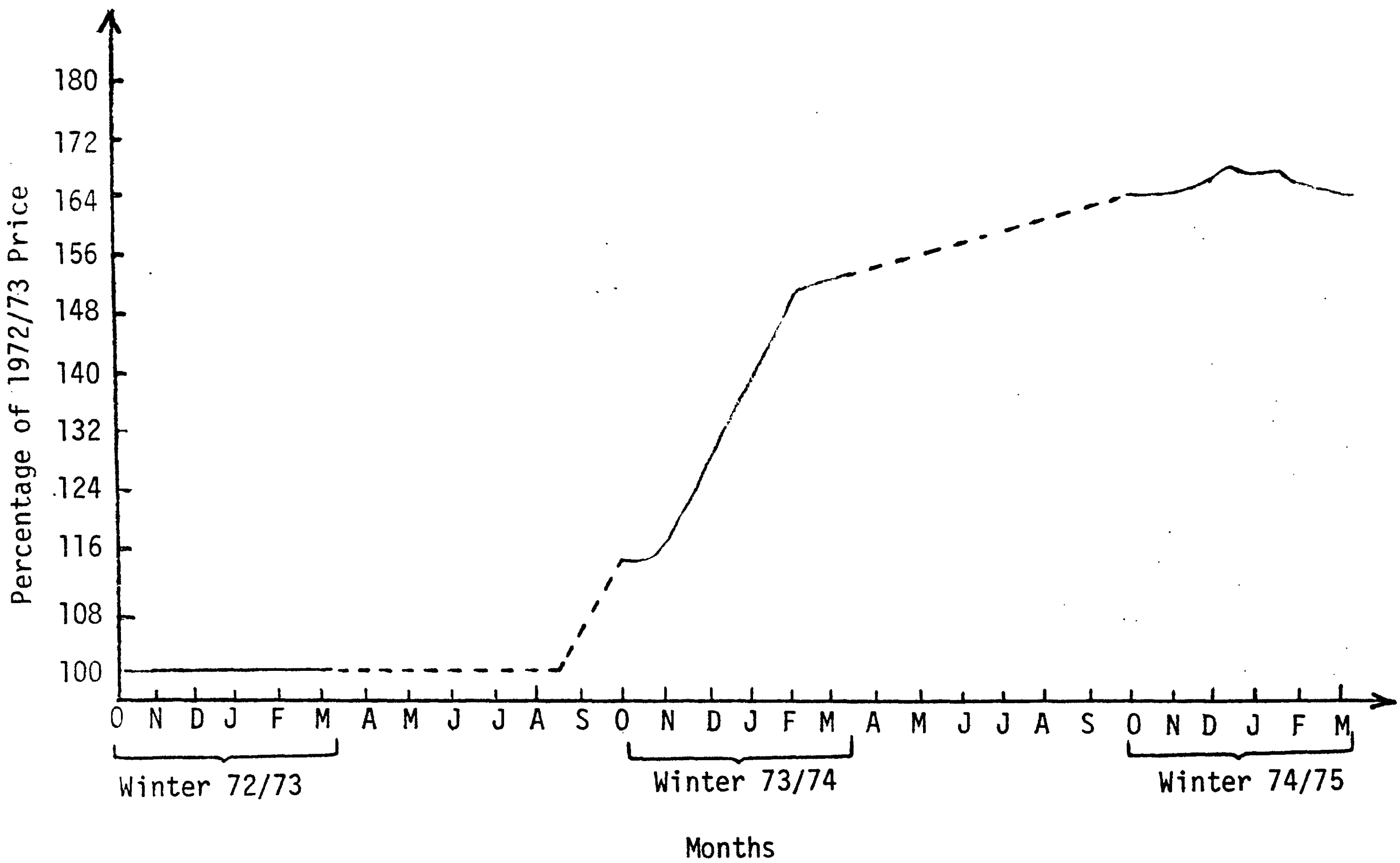


Figure 9: Oil Prices, as a Percentage of 1972 Price, 1972 through 1975

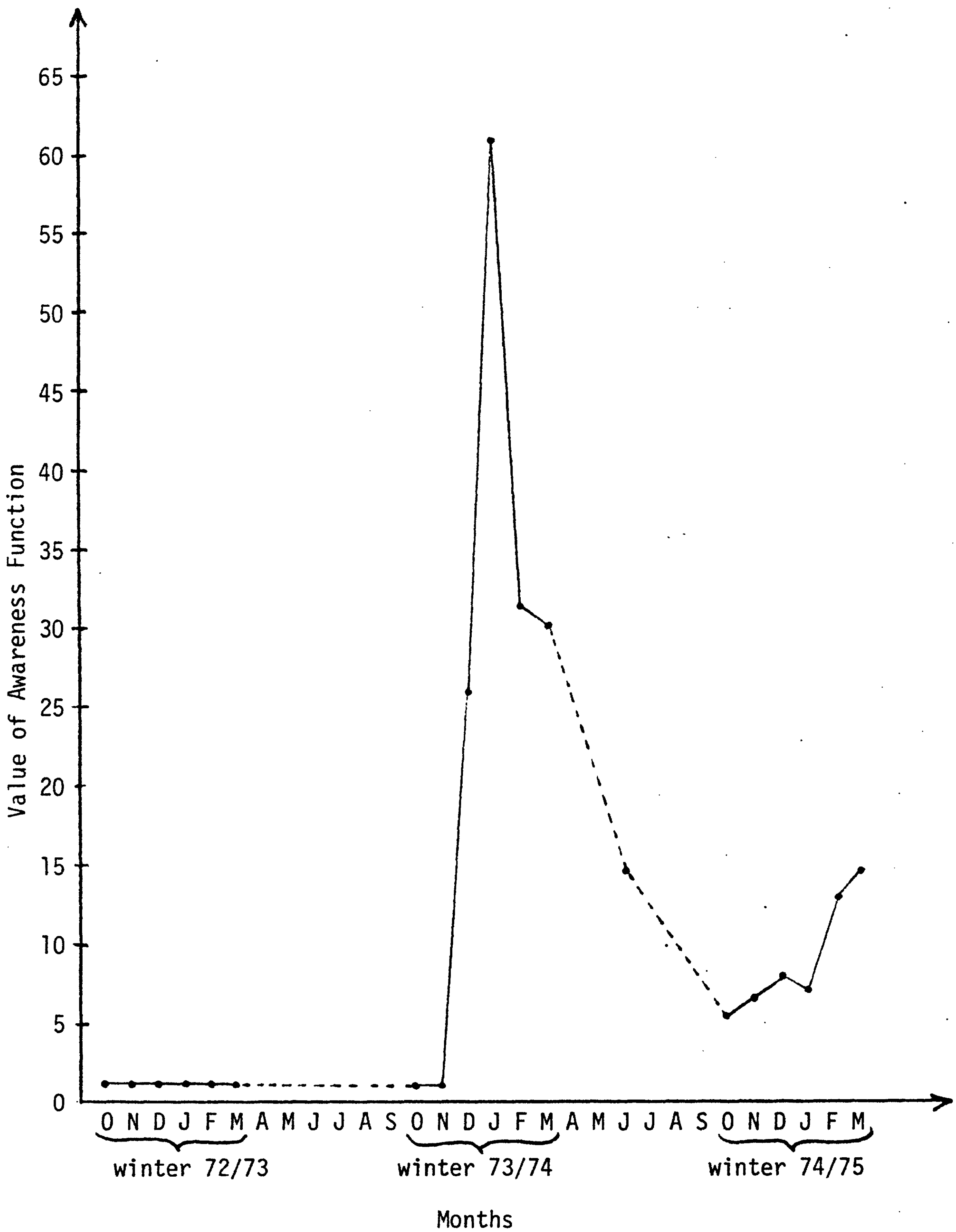


Figure 10: Cumulative Awareness Function, 1972 through 1975

Figure 11 depicts an instance of the model where fuel price data is deflated according to consumer price index for residential fuel as published by the Bureau of Labor Statistics. Further, the cumulated awareness variable was logged and awareness was lagged by two periods.

The price elasticity (evaluated at the sample mean) associated with this model is -.1696. The Durbin Watson statistic indicates no autoregression.

A_i	Term (X_i)	Coefficient	Standard Error	T-Statistic
A_1	Constant	0.28394	0.00262	108.47624
A_2	Deflated Price	-0.00127	0.00009	-14.53209
A_3	Cummulative Awareness	-0.00032	0.00003	-12.58638

Number of observations	12
F-statistic	378.90827
Standard error	0.00131
r^2-statistic	0.98826
Durbin Watson statistic	2.95726

Figure 11: CPD as a Function of Deflated Price and Log of Cummulative Awareness

6.1 Other Instances of the Model

We feel that the model depicted in Figure 11 is most representative. However, several other variations of that model were attempted, all of them yielding to a price elasticity between -.11 to -.18. The lowest price elasticity (and the lowest r^2) occurred using average price instead of deflated price. Hence, it seems that even in the short run there is some evidence to support the hypothesis that consumers perceive and react to price reduction as a consequence of inflation. However, more data over a longer period would further test this hypothesis.

Adding weather as a linear term, as a quadratic term, and as a logarithmic term to the righthand side failed to show any statistical significance of weather on CPD. The fact that weather did not significantly contribute to the explanation of variance in consumption per degree-day does not imply that weather is not important to changes in total consumption. Appendix B shows that weather has a major impact on total consumption, but the measure we use here is consumption <u>per</u> degree-day, which was chosen to normalize for the effects of weather.

A variation of the model relating consumption to price, cumulated awareness, and weather yields elasticities of -.18, -.18, and -.17 where weather was treated as a linear term, a quadriatic term, and a log term respectively. Other variations using different decay factors in calculating cumulated awareness and different lagging periods were also tested, but none yielded better results. All variations of the model showed remarkable stability, and awareness was always statistically significant.

6.2 Other Studies

An independent study was done by the Federal Energy Administration (PIES model)[29] which resulted in a price elasticity of -.466 for residential demand of heating oil in New England. Empirical evidence, however, tends to support the validity of the model of Figure 11 (yielding an elasticity of -.17), since, as seen in Figure 8, a 12 percent curtailment in consumption was realized in the 1973/74 heating season over the 1972/73 seasons during which time a 50 percent increase in price occurred. Another study in the state of Indiana reports a 14 percent reduction in fuel usage over the same period[30] where similar price increases were experienced. In addition, it is most probable that price alone does not determine consumption levels. Rather, a portion of the savings is probably due to consumer awareness of an energy crisis.

The long-term effects of price changes and fluctuations in awareness deserve more scrutiny. However, it can be at least suggested from these results that increases in price cause consumers to conserve energy and also that consumer awareness plays an important role in consumer conservation efforts. Further study and experiments should be carried out to determine the behavioral changes and home improvements that resulted in the conservation of energy displayed in 1973/74. In particular, these changes should be identified and quantified to direct training and awareness programs.

---

[29] Project Conserve Implementation Plan as Submitted by the Massachusetts Energy Policy Office to FEA, April 1976.

[30] Op cit., Peck & Doering.

7. MODEL OF CONSERVATION ACTIONS

The third objective of this study was to quantify and analyze the low-level mechanisms that were enacted, as a result of fluctuations in price and shortage awareness by consumers in their attempts to affect the desired changes in consumption. A corollary objective was to determine whether different consumer groups reacted differently in their efforts to conserve fuel.

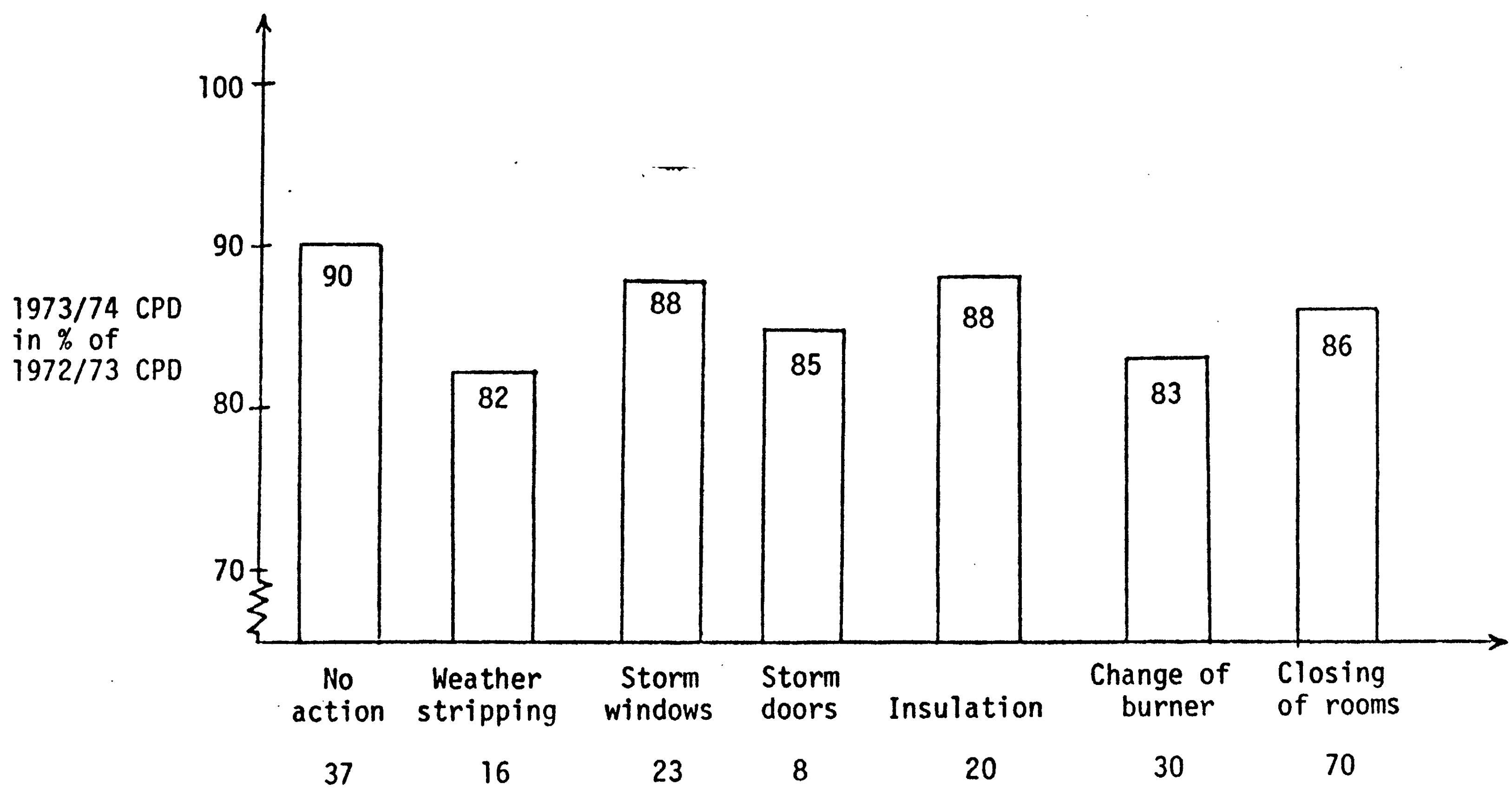
Figure 3 depicts the distribution of CPD for the subsample of '668' single-family homes in suburban Boston, Massachusetts. The mean CPD for 1972/73 was 0.204 gallons per degree-day; for 1973/74, 0.179; and for 1974/75, 0.189. From the means, standard deviations, and sample size given in Figure 8, it follows that the shifts in CPD are highly significant. In an effort to explain this shift, we discuss here our findings on various conservation efforts, e.g., home improvements, thermostat resetting, inventory policy, adding storm windows, and changes in oil burners.

This study indicates that consumer behavior can have an important effect on consumption. For example, between 1972/73 and 1973/74, our sample reduced their consumption per degree-day by 12%. However, the study seems to indicate that this behavior was temporary. That is, in 1974/75 when the "energy crisis" had passed, consumption per degree-day increased by 6 percent from 1973/74.

We can partially account for this savings by reported behavioral changes (from responses to the questionnaire). The change reported most frequently was temperature resetting. That is, 80% reported lowering daytime temperature settings and 72% reported lowering nighttime settings. However, the number of people who reported home improving actions is comparatively small (e.g., under 15% reported any physical changes). The number of households with the subsample of '668' where home improvements were taken in 1973and 1974 is given in Figures 12a and 12b.

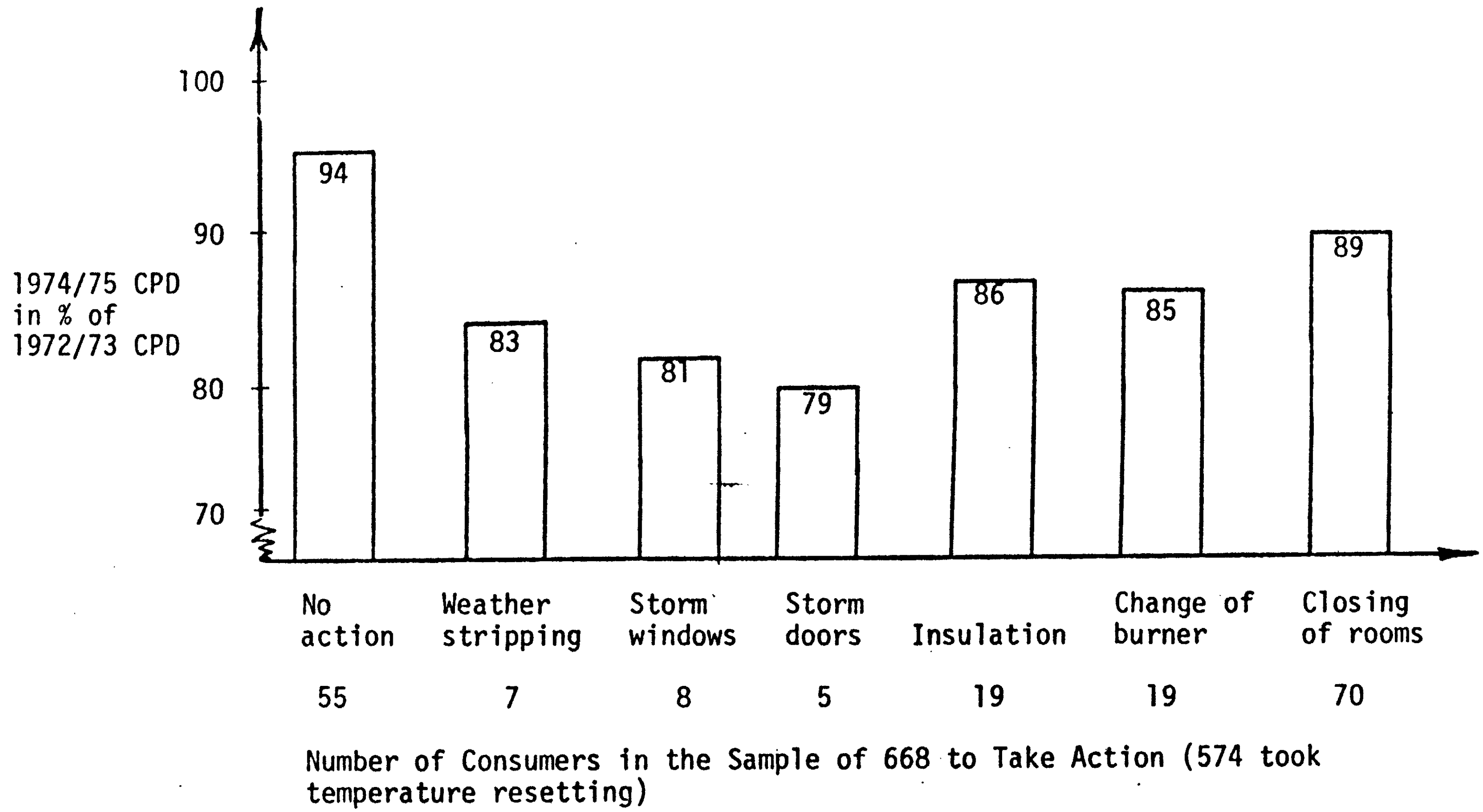
The raw data seems to indicate savings from certain home improvements. Figure 12 indicates, for example, an average CPD reduction of 15% between 1972/73 and 1974/75 for those nineteen customers who installed new oil burners in the summer of 1974. However, not all of that savings can be attributed to the effect of burner replacement. Figure 12 also reveals that consumption went down even in those homes that did not report any explicit behavioral or structural changes. Hence, it appears that consumers' awareness of shortages invokes subtle changes (perhaps front door not left open, closing windows, etc.), which are less easily monitored. Continuous monitoring efforts would be needed to more completely and accurately account for the more subtle changes in behavior. Such efforts were not a part of this study.

Figure 12a: Home Improvements and Actions Other Than Temperature Settings Taken in Summer of 1973 as They Affected Consumption in the Winter of 1973/74



Number of Consumers in the Sample of 668 to Take Action (approximately 600 took temperature resetting)

Figure 12b: Home Improvements and Actions Other Than Temperature Settings Taken in Summer of 1974 as They Affected Consumption in the Winter of 1974/75.



Linear models were used to estimate the influence of consumer characteristics and of specific conservation actions on CPD between heating seasons:

$$\text{MODEL III:} \qquad \Delta CPD_y = B_1 + \sum_{i=2}^{n} B_i Z_i$$

where the dependent variable, ΔCPD_y is the fraction of the 1972/73 consumption per degree-day that was consumed in either y=1973/74 or y=1974/75 heating season and Z_i, i=2, ... are variables that characterize conservation efforts or consumer characteristics. B_i, i = 1, 2, ... are coefficients that are estimated by the ordinary least squares method. Data for ΔCPD_y is based on a subsample of the 668 consumers who provided data on the specific variables used in the regression formula. The conclusion of the bias analysis is that the subset '668' is slightly more energy-conscious than the populace as a whole, and subsets of '668' do not introduce further biases.

A potential problem, resulting from its additive nature, is that Model III would be inadequate if large savings were to result from specific factors and if a number of conservation measures were taken simultaneously by the sample consumers. When the number of simultaneous measures enacted by a consumer is small (i.e., 2 or 3), and as long as the savings per conservation measure is small (i.e., ≈10%), the error that results from the additive model is not significant.

The independent variables were chosen to represent behavioral changes (e.g., temperature resetting), physical changes (e.g., addition of storm windows), and home characteristics (e.g., all the variables of Model I, such as age of homes, income level).

Tests showed that the variables that were significant in Model I are not significant here. Data used in these variables was concluded (Appendix B) to be typical of suburban homes. Thus this model shows no statistical correspondence between these variables and change in CPD, and therefore supports the conclusion that groups within the sample (e.g., sample divided up by income levels, house size, etc.) react in the same manner. These results were further confirmed by an analysis of variance of consumption change for consumer groups with differing consumer characteristics. The independent variables Z_2 through Z_7 were used to denote changes in behavioral or physical characteristics.

Figure 13 reports the results of two regressions with six independent variables:

Z_2 weather stripping, four possible values according to the fraction of the house so equipped

Z_3 installation of storm windows, four possible values according to the fraction of the house so equipped

Z_4 installation of storm doors, four possible values according to the fraction of the house so equipped

Z_5 installation of extra insulation, four possible values according to the fraction of house so equipped

Z_6 change of oil burner, 0 if not present, 4 if present

Z_7 number of degrees the temperature was reset from 1972/73

7.1 Results of Model

All variables have the sign expected. This model confirms the statistically significant impact of temperature resetting[31] (Figure 13), change in oil burner, and addition of insultation upon change in consumption. While Figure 13 represents our best point estimate of the effect of adding weather stripping, storm doors, and storm windows, the data does not present enough variation to statistically confirm our belief in the reduction effect these factors have. Again, this is probably due to a lack of variation within the sample, under 3% of our sample added weather stripping in the two years preceding 1974/75, under 4% added storm windows, and under 2% added storm doors.

The model was run for each of the years with similar results and a similar low r^2 statistic.

The most interesting point to note is that the model explains only a small percentage of the change in CPD_y. That is, while some of the factors included in the model are statistically significant, they explain only 6% of the variance in ΔCPD_y (note r^2 statistics in Figure 11). Other data that we did not ask for and that may be difficult to capture in a questionnaire may be the dominating factors. These may include such behavioral changes as: fewer airings of rooms, sleeping with closed windows, reminding children to shut doors, etc. As these yet-to-be determined behavioral changes may account for the largest part of the 12% savings in consumption between heating seasons 1972/73 and 1973/74, it would seem most promising to initiate further studies to determine what these are and quantify their effect.

---

[31] D.A. Pilati, "Energy Conservation Potential of Winter Thermostat Reductions and Night Setback," Oak Ridge National Laboratory, Tennessee, ORNL-NSF-EP-80, February, 1975.

B_i	Term (X_i)	Coefficient	Standard Error	T-Statistic
B_1	constant	0.87114	0.01137	76.64895
B_2	weather stripping	-0.03623	0.01502	-2.41195
B_3	storm windows	-0.00148	0.01091	-0.13527
B_4	storm doors	-0.02485	0.01433	-1.73385
B_5	extra insulation	-0.01091	0.01186	-0.92042
B_6	change of oil burner	-0.01872	0.00643	-2.91109
B_7	# degrees reset	0.01228	0.00336	3.65933

No of observations	589
F-statistic	6.13578
Standard error	0.17157
r^2	0.05949

Figure 13: Fraction of 72/73 ΔCPD_y consumed in 74/75, taking into account efforts in effect during 74/75, including home improvements made in 1973. Results of regression of CPD change with variables representing conservation efforts.

One of the variables that was recorded and tested with this model that shows an increased level of statistical significance for the winter of 1974/75 as compared with the 1973/74 winter is the installation of insulation. The increased significance of these coefficients could be explained by noting that whereas conservation efforts were intensive for everyone in 1973/74, only consumers who took home improvement actions maintained a level of conservation awareness in 1974/75. Hence, these people continued to successfully conserve energy even after the crisis had ended.

7.2 Cost Effectiveness of Adding a New Oil Burner[32]

In this section Models I and III are used to calculate the cost effectiveness of adding a new oil burner for a household with specific characteristics. Consumers who change oil burners generally upgraded to the Scott Enerjet, manufactured by Beckett Corporation. The new burner is designed to provide a finer mixture of air and oil for more efficient combustion. The newer model uses a 3450 RPM fan to cycle air through the burner where burners made twenty years ago use a slower, less efficient, 1725 RPM fan. In addition, the new burner employs a more efficient diffuser that allows for a better flame retention and thus there is a more complete breakdown of the fuel and more efficient combustion.

From Model III we note that the effect of change of oil burner corresponds to a reduction of 8.99% in CPD (calculated while holding all other variables constant. We can use this percent to calculate total savings for a specific

---

[32]We thank Jerry Hausman for his suggestion to include such a discussion.

house using results from Model I. (Note in the analysis of Model III we supported the hypothesis that different groups of homeowners save the same percentage, i.e., variables such as income, numbers of rooms, etc., were statistically insignificant in Model III. Using Model I the total cost to a household can be calculated for a particular set of household characteristics. Specifically:

$$\text{Cost of fuel for heating season} = [CPD_y] \times [\text{total degree-days in year}] \times [\text{average price per gallon}]$$

$$= [.31647] \times [4250] \times [.41] = \$551.45$$

where CPD_y is calculated using the equation generated in Model I for a ten-room house with a family income of over \$20,000 year, and all other variables set at their mean value. Based on this total cost, changing an oil burner would save such a homeowner \$49.58 that year.

Figure 14 depicts this savings. Also shown is the opportunity costs of capital invested in an oil burner at an alternative interest rate of 5%. As the r^2 statistic for this model is low, even though the t-statistic for this variable is high, there is a confidence associated with this analysis.

The 95% confidence interval (as calculated by

$$\hat{\beta} - (t_{.025})\, \underbrace{S_{\beta}}_{\text{standard error of coefficient}} < \hat{\beta} < \hat{\beta} + (t_{.025})\, S_{\beta}$$

for the coefficient of (B_6) results in a savings range of \$49.58 $\pm$ \$33.37.[33] The shaded areas of Figure 14 depicts this band.

---

[33]Wonnacott and Wonnacott, Econometrics, John Wiley, New York, 1969.

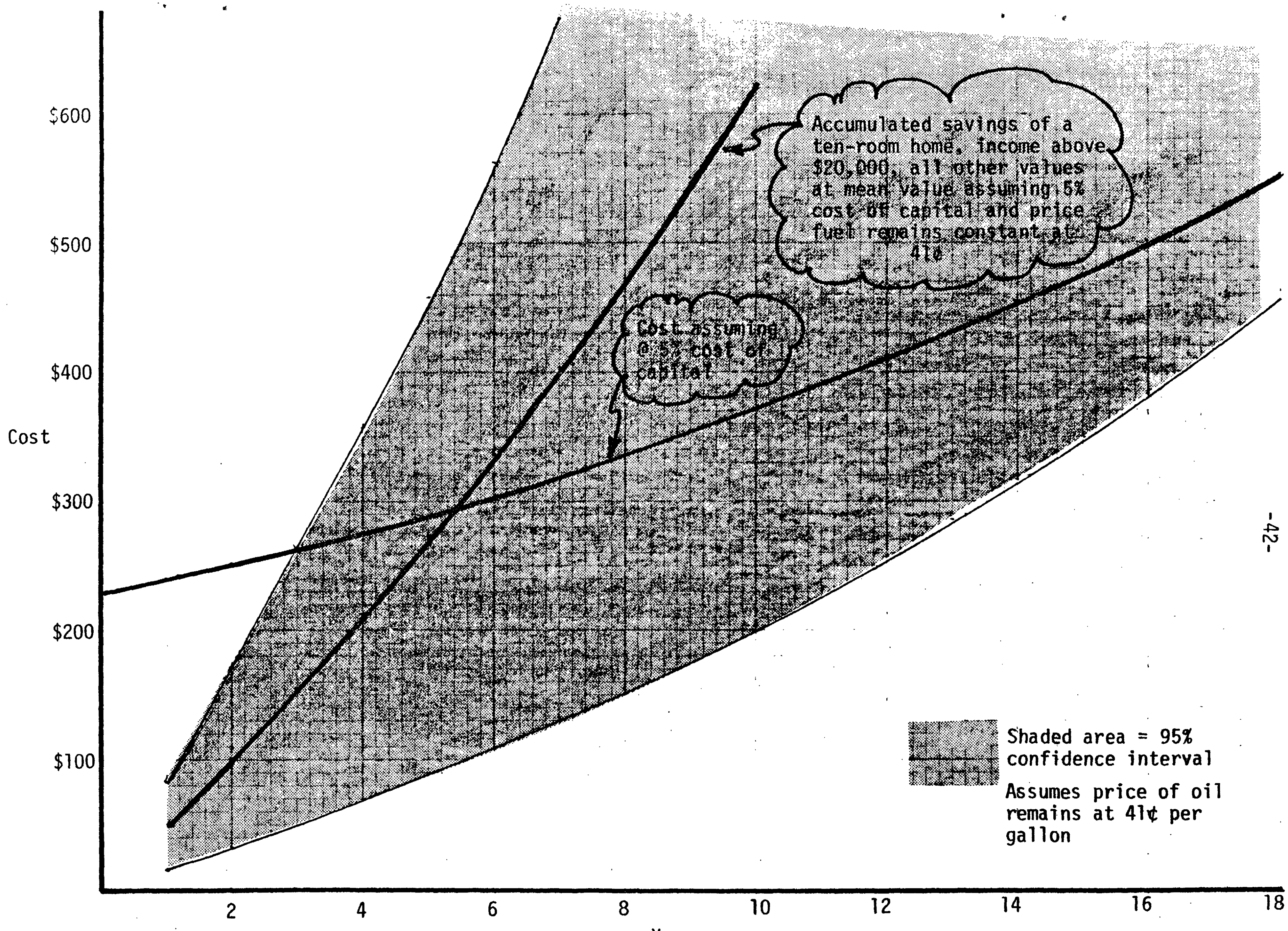


Figure 14: Cost vs. Savings Associated with Change in Oil Burner

8. SUMMARY

The availability of an unusual data series, namely, consumption data on individual homes and corresponding data on characteristics of those homes and occupants has availed us of the opportunities:

(1) To study the factors that are responsible for consumption differences between households. The study indicates that there is a strong correlation between family income, number of rooms, and presence of storm windows on consumption per degree-day. The model indicates that presence of storm doors corresponds to a 4% reduction in consumption, and presence of storm windows throughout a home corresponds to a 11% reduction in consumption over a home without any storm windows.

(2) To examine the effect of price and awareness on residential heating demand in New England. The study shows that corresponding to increased prices and awareness during the heating seasons 1973/74 over 1972/73, there was a savings in consumption of some 12%. Using a time series regression model, the price elasticity was found to be -.17. As the magnitude of this price elasticity seems to be lower than that of other studies, yet agrees more closely with the observed results of our study and of other studies, further investigation is encouraged. The awareness variable was shown to be significant. Yet further work is also needed to determine what factors the populace should be made aware of to encourage energy conservation.

(3) To analyze the behavioral or physical changes that accomplished this savings (these changes were determined to be in response to

fluctuations in price and awareness). Our study indicates that the majority of these savings do not correspond with changes in physical characteristics of homes but in some (yet-to-be-determined) behavioral changes. Temperature resetting, changing oil burners, and adding insulation were found to be significant. The study indicates the cost recovery period of changing oil burners is such that for large homes this seems to be justifiable. For example, for a ten-room home with occupant income greater than $20,000 and all other variables constant, the cost recovery period is approximately five years.

APPENDIX

THE SCOTTLINE

DECEMBER, 1974

Ecomemo . . .

Tom Scott

Home Heating Questionnaire

We are cooperating with the Energy Laboratory of M.I.T. in a study to evaluate residential heating patterns. The study is being conducted in order to better predict our region's continued need for heating oil.

What measures did you take last winter to conserve home heating oil? What steps will you take this winter to heat your home efficiently?

We would appreciate it if you would take a few minutes to fill in the following questionnaire and return it to us with the statement top in the envelope enclosed. Your answers will be correlated with oil consumption patterns of last year, and be kept confidential.

If you would like to make any further comments, because the questionnaire does not fully cover your heating situation, please feel free to do so.

Once the facts are in, we'll report to you in a future issue on the effectiveness of customer conservation measures. This information could well help you reduce your heating bill.

About when was this house originally built? 1940 (year)

On how many sides is the house joined (attached) to other buildings? Ø
(0, 1 or 2 sides)

Is this a one-family or a multi-family house? ONE (One or Multi)

Does the house have a basement or crawl space? YES (Yes or No)

If so, approximately how high above ground does the basement wall reach?
4FT. (feet)

Does the house have an unheated attic? YES (Yes or No)
An attic is considered an unheated area. If it is a heated area, consider your attic as an additional story.

How many stories, not including any basement story, does the house have? ONE

Of what materials are the outside walls? ASBESTOS SHINGLE

Is the house insulated (e.g. by an insulation material like glass fibre)?

walls YES (Yes or No)

ceiling YES

floors NO

attic (if there is one) NO

Your answers to the following questions will help to determine how well home improvement and other physical fuel saving measures work in the average house.

Have any of the following measures been taken?	(Yes or No)	If yes, which part of the house is equipped (whole, ¾, ½, or ¼)?	If the installation took place within the last two years, state: month and year
weather stripping	YES	WHOLE	SUMMER 1974
installation of storm windows	YES	WHOLE	SUMMER 1974
installation of storm doors	YES	WHOLE	FALL 1973
installation of extra insulation	NO		
change of burner	NO		
other technical measures:			

Is the oil you purchase used to provide heat for one or more families in separate living quarters? ONE (One or Several)

If it is one family, please continue with page 3.

If it is several families, please turn to page 4.

QUESTIONS FOR SINGLE-FAMILY ACCOUNTS

How many rooms of your home are heated by oil? 5
(Do not count, bathrooms, porches, foyers, halls or half-rooms.)

How many persons normally live in this home? 3

Was your home unoccupied for any extended period (5 days or more) during the months of October, 1973, through March, 1974 (because of a winter vacation, for example)? ~~YES~~ (Yes or No)

Temperature resetting and the number of rooms kept unheated permanently affect the amount of fuel required. Estimate the temperature setting for your home, and indicate how many rooms of the ones identified above were kept unheated:

During the winter of 1972-73 (before the fuel crisis);
Example: 70° during the day, 60° during the night.

days: 72 degrees nights: 65 degrees

Number of rooms kept unheated permanently NONE

Winter, 1973-74 (during the fuel crisis):

days: 68 degrees nights: 65 degrees

Number of rooms kept unheated permanently NONE

Winter, 1973-74 (after it became apparent that adequate fuel was available):

days: 68 degrees nights: 65 degrees

Number of rooms kept unheated permanently NONE

How do you intend to heat your home this winter?

days: 68 degrees nights: 65 degrees

Number of rooms kept unheated permanently NONE

As a last but important aspect, we would like to study to what extent the family income determines the reaction to rising energy cost. Pleast keep in mind that your individual answers will be kept confidential. If, nevertheless, this question offends your sense of privacy, please skip it. In which of the following brackets did the combined income (before taxes) of everyone living in your home fall in 1973? (Please underline the relevant bracket.):

Below 5,000 $/year
5,000 to 10,000 $/year
10,000 to 15,000 $/year
15,000 to 25,000 $/year
25,000 to 50,000 $/year
Above 50,000 $/year

Thank you for taking the time to answer these questions. Please return them to us with your statement top in the envelope enclosed.

APPENDIX B

BIASES IN THE DATA [34]

This section focuses on the biases of data within the sample. We shall show that the price data series is representative of the price changes that occurred throughout New England. We also show that the weather data used is typical of much of New England. However, the data covers years (1972-75), which were substantially warmer than the long-term averages. The major concern for biases are associated with the two data series, consumption and the questionnaire responses. We shall show that the sample is typical of suburban homes, and that the consumption patterns found in the most often used subsample of '668' are slightly more conservation-minded than for the populace as a whole.

Figure B.1 depicts the sets and subsets of consumers connected with this study. The set marked "USA" denotes all single-family dwelling units in the United States. The subset "N.E." denotes those residential dwelling units in New England. The subset "Massachusetts" denotes those in Massachusetts. The subset labelled '18000' denotes the set of consumers who received questionnaires. The subset '2000' denotes those single family households who responded to be questionnaire. The subset '8000' represents those consumers for whom the oil distributor had detailed monthly consumption data.

---

[34] We acknowledge the assistance of Henry D. Jacoby for his advice on this section.

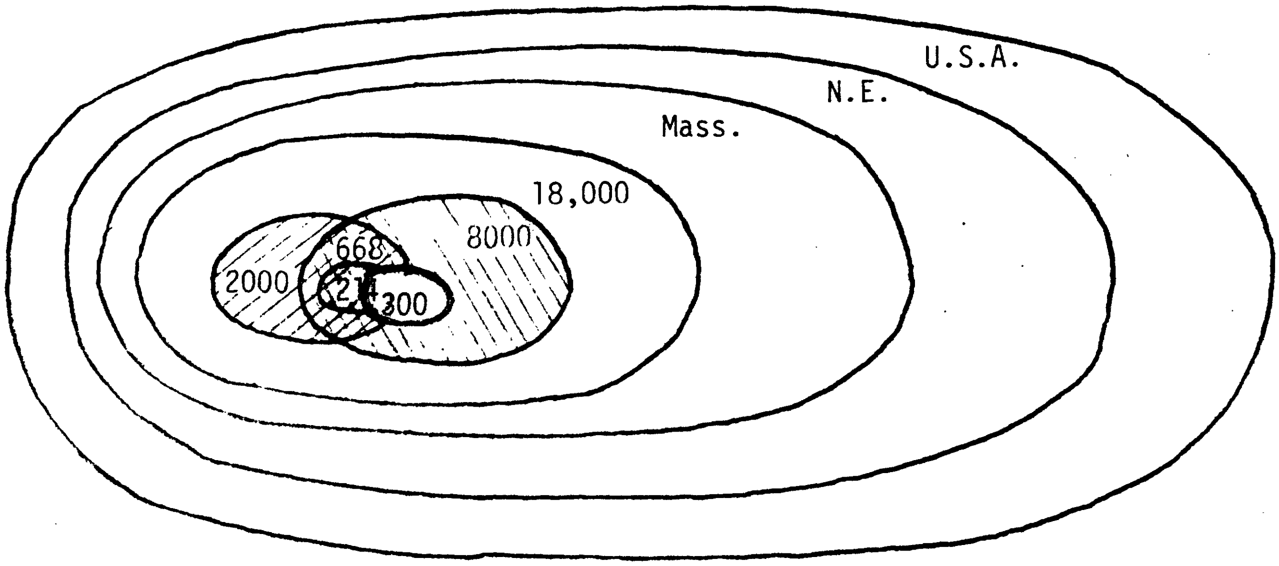


Figure B.1: View of Sample

Section B.1 is concerned with the bias in consumption data, especially in the three subsamples of the 8000 consumers for whom consumption data is available:

a) the 668 respondants for whom consumption data and question-naire responses are available;

b) subsets of the 668 that are used in the consumption Model I and in consumption Model III; and

c) the subset '300' for which frequent oil deliveries were made and which is used in Model II.

Section B.1 is concerned with the possible biases of physical and occupant characteristics within the subsample '668' with respect to the set Massachusetts and USA.

Section B.3 is concerned with bias in price for the oil company with respect to New England.

Section B.4 is concerned with weather biases.

B.1 Consumption Bias of the Sample

Subsets of the consumption data are used to calculate the dependent variable in all three models. With respect to consumption bias of the sample, we present an analysis of four hypotheses:

(1) Individuals who answered the questionnaire are more energy-conscious.

(2) The 668 consumers changed their consumption habits more dramatically during price and shortage changes with a more lasting effect than those who did not respond to the questionnaire.

(3) Consumers who answered a large percentage of the questionnaire do not differ in their energy consciousness from consumers who answered a smaller percentage of questions. In other words, subsets of the '668' based on questions answered do not introduce further bias.

(4) Consumers who receive frequent oil deliveries and are therefore ideal suppliers of information on monthly changes of CPD do not change their consumption more drastically than the average consumer of the '8000' sample.

With respect to hypothesis (1), Figure B.2 denotes the average CPD for set '8000' (all the consumers for whom consumption data was available and who received the questionnaire) and the average CPD for set '668' (consumers who returned the questionnaire). Note the difference in consumption is small (approximately 3%). However, because of the relatively large subset, '668', that difference is statistically significant at a 99% confidence level.

With respect to hypothesis (2), Figure B.3 denotes that percent change in CPD between the heating seasons 1972/73 and 1973/74, as well as the change between 1972/73 and 1974/75. Note that the change between 1972/73 and 1973/74 (the energy crisis) indicates that the 668 sample reduced its consumption somewhat more than the rest of the population, and with regard to the change between 1973/74 and 1974/75, the sample stayed at a lower level of consumption than did the general populace. Hence, the sample consumers remained more energy-conscious longer. These differences are highly significant.

The explanation for these results is that the sample is a more conservation-minded group than the average occupant of a one-family house. This bias implies that our results are more likely to be optimistic with respect to the conservation efforts undertaken by the population as a whole.

Statistic \ Season	1972/73	1973/74	1974/75
Mean CPD of Set '8000' (Standard Deviation)	.2099 (.082)	.1873 (.078)	.1960 (.078)
Mean CPD of Set '668'	.2021	.1793	.1870

Figure B.2: Consumption Bias of Sample '668'

Time Period / Statistic	between 1973/73 and 1973/74	between 1972/73 and 1974/75
Subset '8000' Mean Change in CPD	.8974	.9637
(Standard Deviation)	(.1363)	(.2435)
Subset '668' Mean Change in CPD	.8820	.9278

Figure B.3: Percent Change in Consumption Sample '8000' and Sample '668'

With respect to hypothesis (3), certain questions on the questionnaire met with a larger percentage of non-responses than others. This means that with the introduction of different consumer characteristics into the analysis, the size of the useful subsample varies. The smallest subsample considered was used to compute the correlation matrix of consumer characteristics and consumption data (Appendix C), which is based on 214 consumers who provided data on all 32 correlated items. Hypothesis (3) is tested by comparing the consumption per degree-day of the 668 and the 214 sets.

The differences in the parameters depicted in Figure B.4 is insignificant. This means there is no indication of an additional bias caused by further subsampling through elimination of consumers who did not answer specific questions.

B.2 Physical and Occupant Characteristics Bias of Sample

Using data obtained from the questionnaire and from the 1970 U.S. Census Report, the bias in physical and occupant characteristics of the '668' sample are investigated. Figures B.6, B.7, B.8, and B.9 depict comparisons between the sample and the subsets Massachusetts and U.S. with respect to four physical or occupant characteristics. Our conclusions are that demographically (number of persons per household) the sample is not significantly different from the homes in the nation. In general, the sample shows characteristics typical of suburban single-family homes in which income is higher than the national average, homes are of larger size, and house ages are younger than in the "old" state Mass. but compare well with the U.S. as a whole.

Well over 50% of the American population lives in suburbs (70% in single-family homes). While the sample is weighted toward those characteristics of suburban homes, it has consumers that cover the entire spectrum of income levels, age of homes, and number of rooms and therefore allows us to study the characteristics and behavior of such subgroups.

deliveries provides a great deal of information on the change of consumption within this heating season. Hence, the '300' set is important in calculating useful monthly consumption data.

Figure B.5 depicts the relative change in $\overline{CPD}_y$ between heating seasons for a subsample of the '8000' that receives frequent deliveries and for the '8000' sample. Note that Figure 15 shows only small, insignificant differences for 1973/74. However, the differences between 1974/75 are larger and significant. A possible explanation for this is that since the sample '300' represents consumers who received frequent oil deliveries and since shortage awareness stayed higher during the 1974/75 season, this sample consumed less. Note that this group did increase their consumption in 1974/75 but not by as much as the population as a whole. Hence, we can assume that using this subsample in the price model does not considerably change the size of the resulting price elasticities.

Statistic \ Season	1972/73 to 1973/74	1972/73 to 1974/75
Mean '8000'	.897	.964
(Standard Deviation)	(.136)	(.244)
Mean '300'	.896	.936
(Standard Deviation)	(.121)	(.169)

Figure B.5: Bias of Sample '300'

Statistic \ Season	1972/73	1973/74	1974/75
Mean CPD of Set '668'	.2099	.187	.196
(Standard Deviation)	(.082)	(.078)	(.078)
Mean CPD of Set '214'	.203	.176	.184
(Standard Deviation)	(.078)	(.064)	(.066)

Figure B.4: Sample '668' and Sample '214'

Hypothesis (4) is important as it deals with bias introduced in choosing a subsample of those consumers who have five or more deliveries in all three heating seasons. This subsample is used in the price model where the influence of changes in consumption corresponding to price and awareness fluctuations are examined.

Using our source for consumption data (the oil company delivery records), consumption changes can only be measured whenever a delivery occurs. Consumption in a period is equal to the quantity delivered, where a period is defined as the date between this delivery and the previous one. Hence, those customers who have frequent deliveries provide more reflective data on consumption, since consumption is monitored more often. For example, a customer who receives only one delivery during the heating season provides no information on the change of consumption during the heating season, whereas a customer who recieves six

Income	Massachusetts	U.S.	Our Sample
less than $5000	10	20	10
$5000-10,000	27	33	17
$10,000-15,000	38	27	16
$15,000-25,000	20	16	29
$25,000-50,000	5	4	22
greater than $50,000	1	1	5

Figure B.6: Income Distribution (% of Sample)

Number of Rooms	Massachusetts	U.S.	Our Sample
1 - 4	12	18	4
5	24	29	12
6	29	27	30
7	18	14	21
8	10	7	16
9 - 12 +	7	5	14

Figure B.7: Number of Rooms in Owner Occupied Dwellings (% of Sample)

Year Built	Massachusetts	U.S.	Our Sample
1971-74	--	--	5
1961-70	19	26	19
1951-60	21	25	24
1941-50	8	13	20
Before 1941	52	36	32

Figure B.8: Age of Dwelling (% of Sample)

Number of Persons	Massachusetts	U.S.	Our Sample
1	10	12	14
2	27	30	35
3	17	18	15
4	18	17	19
5	13	11	10
6	8	6	4
7	4	3	2
8 +	3	2	0

Figure B.9: Persons per Dwelling (% of Sample)

B.3 Price Bias

The period 1972 through 1975 was marked by price increases of fuel oil that were larger than any in recent history. From the start to the finish of the heating season of 1973/74 there was a 30% increase in price. During the entire period of 1972 through 1975 there was a 50% increase in price.

Examining the prices of Scott Oil Company and average price data for the entire region, the changes in price are similar and occurred at the same time (.99 correlation with average prices for all of New England). The conclusion is that the price data used is reflective of the entire region.

B.4 Weather Bias

In this section we conclude that;

- The winters 1972 through 1975 are warmer than normal.
- The winters 1972 through 1975 are similar.
- Consumption is highly correlated with weather.
- Weather variations well within a probable range can have a large effect on consumption of heating oil.
- CPD appears to be linear in the range our data covers.
- The weather for the area of our sample does not differ substantially from the rest of New England.

B.4.1 Bias of Warmer Winters

Weather is the single most significant factor affecting consumption. Our data suggests a correlation of over .90. The degree-day distribution for the period December through March in Boston based on a weather history of forty years would suggest a mean of 4497 and a standard deviation of of 286 degree-days. Hence there is a 10 percent probability that the degree-day totals from December through March in Boston will be outside the range of 4025 and 5969. That is, there is a 90 percent probability that the degree-day variation in Boston will be between $\pm$ 10.5 percent from the long-term forty-year avarage of 4497 degree-days. The winter seasons of 1972/73, 1973/74, and 1974/75 were all warmer than average with degree-days totals of 4144, 4282, and 4250 respectively. Thus it should be noted that this study spans a period of warmer than normal winters.

The implications of this bias are twofold.

(1) The policymakers should be aware that any actions taken to conserve fuel may in fact be offset by weather. For example, if 1973/74 had been cold, yet within the 95 percent probability range of weather distribution for the region, homeowners would have consumed 16 percent more energy. This is not to imply that conservation efforts should not be encouraged, but the policymaker that is associated with such efforts should educate his constituencies as to the effect that weather can have so that his conservation efforts are not considered ineffective as a consequence of a cold winter.

(2) Engineering studies suggest that CPD is not a linear relationship between consumption and weather but rather has a quadratics or higher order relationship. The experience of oil companies suggests that smaller CPD goes with colder weather. We have tried to confirm these observations with our data. For example, using a variant of Model II, we related CPD price,

awareness, <u>and</u> weather. Weather was treated as a linear term, as a quadratic term, and as a log term. Using our data, all variations showed weather to be insignificant. Our data may not have sufficient variation to show a non-linear relationship. We do not have consumption data for very cold winters or for very warm winters.

If CPD is not linear, then the elasticity of -.17 may change for significantly colder or warmer winters.

B.4.2 Weather Representative of New England

The monthly degree-day totals used in this study show weather patterns in Boston that are similar to those experienced by the rest of New England. Boston's weather pattern has a 98 percent correlation with that of Burlington, Vermont, and similarly with Hartford, Connecticut. The degree-day totals in different areas of the region differ greatly even though fluctuations in weather are fairly uniform. Mean monthly degree-day totals in Boston differ by 25 percent from that of Burlington and 17 percent from that of Hartford.

APPENDIX C

	ATTACHMENT TO OTHER H.	EXISTENCE OF A BASEMENT	EXISTENCE OF AN UNHEATED ATTIC	NUMBER OF STORIES	INSULATION OF WALLS	INSULATION OF CEILING	INSULATION OF FLOORS	WEATHER STRIPPING	STORM WINDOWS	STORM DOORS	EXTRA INSULATION	CHANGE OF BURNER	NUMBER OF ROOMS	NUMBER OF PERSONS	OCCUPANCY DURING 73/74	72/73 TEMPERATURE SET. DAY	72/73 TEMP. SET. NIGHT	72/73 NO. OF UNHEATED ROOMS	73/74 DURING THE CRISIS TEMPERATURE SET. DAY	TEMP. SET. NIGHT	NO. OF UNHEATED ROOMS	73/74 AFTER THE CRISIS TEMPERATURE SET. DAY	TEMP. SET. NIGHT	NO. OF UNHEATED ROOMS	74/75 TEMPERATURE SET. DAY	74/75 TEMP. SET. NIGHT	74/75 NUMBER OF ROOMS UNHEATED	FAMILY INCOME	CPD IN 72/73	CPD-CHANGE 72/73 TO 73/74	CPD-CHANGE [illegible] TO 74/75
YEAR OF CONSTRUCTION	10	-10	8	-41	35	46	24	2	7	7	5	-16	-22	4	-6	11	13	-15	15	15	-5	5	7	-1	6	7	-1	9	-34	4	-5
ATTACHMENT TO OTHER H.		1	-8	2	5	8	5	9	-5	10	4	0	-2	13	-12	-2	8	-4	-6	3	-9	-1	3	-9	-4	-5	-10	-5	-5	5	0
EXISTENCE OF A BASEMENT			-7	13	-14	-11	1	1	6	-9	-1	-2	-2	-2	1	-7	-15	6	-3	-11	0	0	-9	0	-5	-11	2	-7	6	-14	-14
EXISTENCE OF AN UNHEAT. ATTIC				-20	0	3	2	-2	9	5	2	13	-11	-9	6	6	-13	5	16	-5	3	23	-3	2	15	-5	-1	3	-18	13	18
NUMBER OF STORIES					-24	-35	-6	-3	-11	-14	-7	13	40	19	2	-6	-3	4	-1	-2	12	-7	-3	11	-5	-5	9	6	40	3	-3
INSULATION OF WALLS						51	36	12	-8	0	10	-19	-3	7	-11	5	3	6	5	2	12	-4	-2	9	-6	0	7	17	-15	4	-2
INSULATION OF CEILING							39	10	-2	-1	12	-14	-3	11	-8	0	6	-10	0	8	-2	-2	5	-2	0	7	-4	24	-18	-5	-3
INSULATION OF FLOORS								9	-8	-5	0	-5	13	21	-2	-5	6	-1	-2	7	10	-3	5	8	-5	3	5	21	-4	4	3
WEATHER STRIPPING									20	10	11	-2	1	6	-1	-1	-9	-7	-1	-13	0	-7	-16	0	-11	-21	-4	4	1	11	6
STORM WINDOWS										36	19	8	-16	-4	-1	13	-2	-12	16	-5	-3	3	-9	-4	9	-7	-2	-19	-22	1	-5
STORM DOORS											12	7	-22	-9	9	2	-1	-1	12	-5	-6	9	-4	-2	15	-3	-5	-13	-24	-5	0
EXTRA INSULATION												3	8	9	16	-11	-5	-3	3	-5	6	-2	-7	8	0	-8	8	10	4	3	-1
CHANGE OF BURNER													10	-19	16	3	-5	0	13	-4	5	11	-4	6	12	-3	5	-4	10	10	8
NUMBER OF ROOMS														42	4	-11	1	1	-15	-4	11	-20	-5	7	-14	-6	9	42	59	0	-5
NUMBER OF PERSONS															-18	-17	11	-12	-13	7	-8	-19	3	-10	-18	-1	-10	37	23	-7	-5
OCCUPANCY DURING 73/74																1	-1	14	-4	2	10	0	6	11	4	5	9	15	3	-2	7
72/73 TEMPERATURE SETTING DAY																	37	3	43	17	4	49	16	3	47	20	0	-11	-5	-6	2
72/73 TEMPERATURE SETTING NIGHT																		-13	10	70	-26	13	69	-26	10	62	-19	8	4	-8	-6
72/73 NO. OF UNHEATED ROOMS																			2	-7	73	2	-7	75	-8	-10	75	0	23	-8	-6
73/74 DUR. CRISIS TEMP. SETTING DAY																				32	4	76	24	4	67	23	3	-5	-6	9	4
73/74 DUR. CRISIS TEMP. SET. NIGHT																					-21	23	93	-19	16	83	-17	5	0	-3	-10
73/74 DUR. CRISIS NO. OF UNHEATED ROOMS																						-2	-24	96	-5	-23	87	12	20	-2	-6
73/74 AFTER CRISIS TEMP. SETTING DAY																							31	-1	85	31	-2	-4	-10	6	18
73/74 AFTER CRISIS TEMP. SETTING NIGHT																								-23	22	89	-19	5	-1	-3	-2
73/74 AFTER CRISIS NO. OF UNHEATED ROOMS																									-6	-22	86	9	21	-6	-9
74/75 TEMP. SETTING DAY																										30	-5	-2	-18	2	18
74/75 TEMP. SETTING NIGHT																											-18	5	-5	-10	-4
74/75 NO. OF UNHEATED ROOMS																												10	22	-7	-10
FAMILY INCOME																													37	-8	-1
CPD IN 72/73																														-18	-21
CPD-CHANGE 72/73 TO 73/74																															74

Correlations between Consumer Characteristics (Each Coefficient is Multiplied by 100)