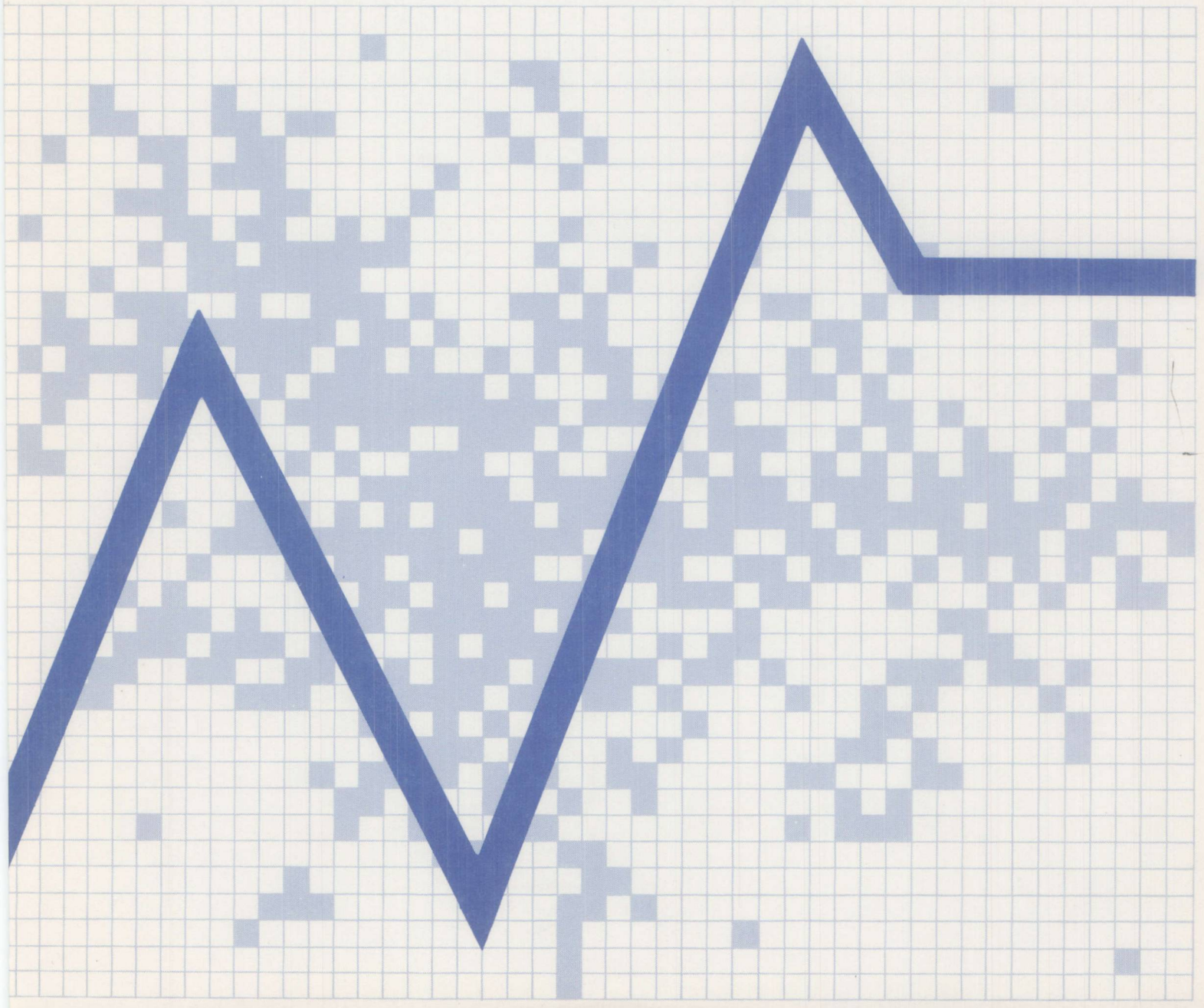
DOE/EIA-0272/S

The National Interim Energy Consumption Survey:

Exploring the Variability in Energy Consumption - A Supplement

October 1981

U.S. Department of Energy
Energy Information Administration
Office of Energy Markets and End Use



This publication is available from the Superintendent of Documents, U.S. Government Printing Office.

Superintendent of Documents
U.S. Government Printing Office
Washington, D.C. 20402

Order Desk: (202) 783-3238

Stock Number: 061-003-00217-0
Price: $3.25

For questions on energy statistics or information on availability of other EIA publications, contact.

National Energy Information Center, EI-20
Forrestal Building
U.S. Department of Energy
Washington, D.C. 20585

(202) 252-8800

For those living and working in the Mountain and Pacific time zones, you may call:

National Energy Information Center Affiliate
Albuquerque, N.M. 87131

(505) 846-2375

DOE/EIA-0272/S
UC-13

The National Interim Energy Consumption Survey:

Exploring the Variability in Energy Consumption - A Supplement

October 1981

U.S. Department of Energy
Energy Information Administration
Office of Energy Markets and End Use
Energy End Use Division
Washington, D.C. 20585

This report was prepared by Robert B. Latta.

Released for printing October 7, 1981.

PREFACE

This report supplements the report entitled, The National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, which presented a preliminary analysis of residential energy consumption models for households living in single-family detached dwellings. In order to fully understand this report, it is necessary to also read the above report. The only fuels for which models were developed were electricity and natural gas. This report expands that analysis to fuel oil, kerosene, and liquid propane gas (LPG). In addition, models for households living in mobile homes, single-family attached dwellings, and units in small and large multi-family buildings are added whenever enough data are available to adequately estimate the terms in the model.

This report and the report it supplements will be of interest to researchers that are studying the possible impact of natural gas deregulation. This represents a large potential audience. The technical style in which this report is written reflects the audience for which it is written.

CONTENTS

TABLES

Page

FIGURES

1. INTRODUCTION

The report, National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, (DOE/EIA-0272), examined the variability in electricity and natural gas consumption for households living in single-family detached units. This report extends the results of the above-mentioned report to fuel oil, kerosene, LPG, and other housing types.

Section 1 summarizes the electricity and natural gas models that are contained in the above-mentioned report. In addition, the methodology that was used to obtain the models is reviewed.

In extending the results, the techniques used depended on the number of households available for a given fuel and housing type. Section 2 describes the amount of data available by fuel and housing type. There is enough data to fit the model for fuel oil use for households living in single-family detached dwellings. Section 3 gives the results for this model and compares it to the electricity and natural gas models. For other fuel and housing types, some adjustments to the model building methodology are necessary due to small sample sizes. Section 4 outlines the adjustments that were used, presents the modified models and discusses the results. Section 5 discusses all of the models and the limitations of their use and interpretation.

2. SUMMARY OF ELECTRICITY AND NATURAL GAS MODELS FOR SINGLE-FAMILY DETACHED DWELLINGS

The electricity and natural gas models are regression models that expressed the predicted electricity and natural gas consumption in MBtu (1000 Btu) as functions of characteristics of the household and housing unit. The models were developed using data from the National Interim Energy Consumption Survey (NIECS), which covers the period April, 1978 to March, 1979.

The details of the NIECS sampling plan are summarized in the report, National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, (DOE/EIA-0272). A discussion on the theory behind the particular form of the models chosen and the choice of the independent variables that were used in the models is also contained in the report. The list of variables that were used in these models as well as those in this report are contained in Appendix A along with a brief description of what they measure.

Only those households living in single-family detached dwellings, having good utility data and for which a personal interview was obtained were used in the modeling procedure.

We felt that the dynamics of space-heating and cooling may differ appreciably between housing types. Consequently, we initially restricted consideration to households living in single-family detached dwellings because it was the dominant housing type.

Various techniques were used to convert electricity and natural gas billing period consumption data to an annual figure. See National Interim Energy Consumption Survey Report on Methodology Part I. Household and Utility Company Surveys, Response Analysis Corporation, Princeton, New Jersey for details. The cut-off point between the technique used for an intermediate amount of consumption data and the one used for "nearly" a full year of consumption data was 330 days. Therefore, we defined "good utility data" for electricity and natural gas as those annual consumption figures that were based on at least 330 days of consumption data.

The mail questionnaire did not contain all of the questions that we asked during a personal interview. Most of the household and housing unit characteristics were imputed for households where a mail questionnaire was used. For this reason, we decided not to use households for which we did not have a personal interview.

For households where the household and housing unit data were obtained using a personal interview, there were still cases of nonresponses to various questions. If the nonresponse rate for a particular question was small, then the missing values were imputed. See National Interim Energy Consumption Survey Report

on Methodology, Part I. Household and Utility Company Surveys, Response Analysis Corporation, Princeton, New Jersey for details. This is not true for questions with a large number of nonresponses. The number of square feet of floor space in the residence and various measures of insulation fall into this last category.

We decided to treat the presence of imputed values, other than consumption figures, as part of the error term. When we did the analysis we did not have a data set available that indicated which responses were imputed. Taking the presence of imputed values into account would have delayed the analysis. There was some evidence that a large number of households had at least one response that was imputed. Removing these housholds from the data set may have greatly reduced the sample size available for building regression models.

The models were obtained using a form of robust regression that required several steps. The steps consisted of fitting least square solutions to the model and an outlier rejection procedure. The general linear model procedure of the SAS statistical package was used throughout.

In the first step of each analysis, parameters were fit to the data using ordinary least squares. For each succeeding step up to a total of seven, new parameters were fit using a weighted least squares procedure with the weights equal to the reciprocal of the estimated consumption from the previous step. The weights reflect the observation that there is a larger variance in energy use among households with high potential use than among those with low potential use.

At the third iteration, we began systematically eliminating outliers before refitting the models. The process of elimination and refitting was continued until no further outliers were removed, or until the seventh iteration was reached. For more details, see National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, (DOE/EIA-0272).

The same procedure was used in obtaining the additional models presented in this report, but in all except one case, the only independent variables that were used were based on models from other cases. This will be discussed further in the next two sections.

Tables 1 and 2 give the final models obtained for natural gas and electricity consumption for households living in single-family detached dwellings.

The income variable used in the electricity consumption model is measured on an ordinal scale. This scale is given in Appendix A. It gives the category in which the respondent places the household's income for 1977. The end points of the categories would have to be changed if the model is to be used with current income data.

The code for NGASNDX, the gas appliance index, and for KYHSBREC, year house built, are also given in Appendix A.

TABLE 1. WEIGHTED LEAST SQUARE SOLUTION FOR NATURAL GAS MODEL (USING NCNGYRB AS THE DEPENDENT VARIABLE).

PARAMETER	ESTIMATE	T-STATISTIC FOR H0: PARAMETER=0	ATTAINED 2-SIDED LEVEL OF SIGNIFICANCE	STD ERROR OF ESTIMATE
INTERCEPT	7,885.86117328 *	2.97	0.0030	2,653.97724959
NGASNDX	478.57038686	3.28	0.0011	145.88367773
NHSLDMEM*HGASWHT	8,183.12797801	12.97	0.0001	631.08680052
HSBNMGSH	39,380.14733876	5.26	0.0001	7,488.38074856
KMHTEQRC				
RADIATORS OR HOT WATER PIPES (1)	35,976.62454035 *	6.10	0.0001	5,899.72507242
CENTRAL FORCED AIR (2)	21,177.30479874 *	5.03	0.0001	4,206.54981219
OTHER (3)	7,946.83589671 *	2.14	0.0323	3,709.19030002
NO GAS MAIN HEATING (4)	0.00000000 *	**	**	**
NHEATDD*HGASMHT*NBATHRMS	4.35890180	7.91	0.0001	0.55073749
NHEATDD*HGASMHT*NAGE01	0.05823754	5.01	0.0001	0.01163191
NHEATDD*HGASMHT*KYHSBREC	-1.12679365	-8.12	0.0001	0.13882925
NHEATDD*HGASMHT*NROOMS	1.14812215	6.45	0.0001	0.17801440
NHEATDD*HGASMHT*NDRSAWS	0.36930515	7.11	0.0001	0.05192832
NHEATDD*HGASMHT*NSDRSAWS	-0.19222837	-4.62	0.0001	0.04156752
NCOOLDD*NRMGASAC	3.97906768	4.55	0.0001	0.87531020

* THE SOLUTION IS NOT UNIQUE BECAUSE THE DESIGN MATRIX IS SINGULAR.
** NO TEST IS GIVEN BECAUSE THE ESTIMATE WAS ARBITRARILY SET EQUAL TO ZERO.

TABLE 2. WEIGHTED LEAST SQUARE SOLUTION FOR ELECTRICITY MODEL (USING NCELYRB AS THE DEPENDENT VARIABLE).

PARAMETER	ESTIMATE	T-STATISTIC FOR H0: PARAMETER=0	ATTAINED 2-SIDED LEVEL OF SIGNIFICANCE	STD ERROR OF ESTIMATE
INTERCEPT	1,838.90033348	1.32	0.1874	1,394.29804463
NELCKDV	1,542.91063357	6.06	0.0001	254.42416974
KINCOME	249.49782225	2.47	0.0135	100.89977992
NAGE01	-46.19145874	-3.01	0.0026	15.32376895
NHSLDMEM	1,800.26754569	10.21	0.0001	176.30579562
NHSLDMEM*HELWHT	3,567.83543487	16.11	0.0001	221.41854047
HELWHT*HELDISHW	5,010.04258170	4.08	0.0001	1,227.15715352
NELFRIG	3,176.97311550	5.45	0.0001	582.64498142
HELDISHW	3,246.81149591	5.14	0.0001	631.57851582
HELCLSDY	3,827.05159181	7.31	0.0001	523.81387074
HSPFDFRZ	4,456.46196110	9.69	0.0001	460.08932694
HAUTOWSH	2,441.14974480	4.22	0.0001	577.80748929
NTOTAL	108.33553652	3.57	0.0004	30.34494681
HSBNMELH	3,282.23184877	4.23	0.0001	776.69185130
NHEATDD*HELMHT*HELHTPUM	-2.56940122	-2.72	0.0066	0.94493510
NHEATDD*HELMHT*NAGE01	0.04376684	4.27	0.0001	0.01025377
NHEATDD*HELMHT*NBATHRMS	2.16693210	5.19	0.0001	0.41768602
NHEATDD*HELMHT*NSDRSAWS	-0.11337950	-3.04	0.0024	0.03727163
NHEATDD*HELMHT*NDRSAWS	0.31150668	6.37	0.0001	0.04890592
NCOOLDD*NRMELRAC	0.69793901	4.04	0.0001	0.17282322
NCOOLDD*NRMELRAC*KINCOME	0.07027372	2.63	0.0087	0.02675765
NCOOLDD*NRMELCAC*KINCOME	0.12566062	5.81	0.0001	0.02162087
NCOOLDD*NRMELCAC	0.41258501	2.36	0.0184	0.17482890

3. AVAILABILITY OF DATA FOR EXTENDING MODELS

The number of households that use electricity is displayed in Table 3, natural gas in Table 4, fuel oil in Table 5, kerosene in Table 6, and LPG in Table 7. For each housing type, the tables give the number of households in the sample that use each fuel, the number of households for which we have good data, the number of households whose main heating equipment uses the fuel, and the number of households with good data whose main heating equipment uses the fuel.

For electricity and natural gas, "good data" implies the yearly consumption figure was based on at least 330 days of data and the household participated in a personal interview. For fuel oil, kerosene, and LPG, "good data" implies the supplier provided actual fuel purchase data from which an annual consumption amount could be obtained and the household participated in a personal interview.

In order for us to have "good data" for any fuel, the consumption for a particular household must be measured separately. Therefore, we should expect that few households living in large buildings will have "good data."

The consumption and cost figures for fuel oil and kerosene were combined on the NIECS data set. Only 882 of the households in the NIECS data set indicated that they used fuel oil or kerosene. Three of the 882 households did not list fuel oil or kerosene as the main space-heating, secondary space-heating, or water-heating fuel. One of the households listed fuel oil as the main space-heating fuel and kerosene as the secondary space-heating fuel. For these four households, it is not clear which of the two fuels was actually consumed. For this reason, these four households were not classified as having good data for fuel oil or kerosene. They were arbitrarily classified as having consumed some fuel oil but not classified as having consumed some kerosene.

All households with good data paid their own utility bills. As a result, the models that were developed apply only to households that pay for their energy use directly. The energy consumption was imputed for households whose utility bills were part of the rent. Therefore, it would be difficult to make a comparison between the energy use of households that did and did not pay their own bills using the NIECS data set.

The following facts can be gleaned from Tables 3, 4, 5, 6, and 7.

- Most households live in single-family detached units.
- The total number of households that use electricity is 4,080. This is one less than the number of households in the data file.
- The dominant heating fuel is natural gas.

Table 3. Total Households in the Sample that Use Electricity, Have Good Data, Use Electricity for Main Heating Fuel, and Use Electricity for Main Heating Fuel and Have Good Data (Classified by Housing Type).

Housing Type	Total That Use Electricity	Total That Have Good Electricity Data	Total That Use Electricity for Main Heating Fuel	Total That Use Electricity for Main Heating Fuel and Have Good Electricity Data
Unknown	11	2	1	0
Mobile Home (KTYPLVQT=1)	270	177	70	55
Single-Family Detached Unit (KTYPLVQT=2)	2,681	2,147	368	301
Single-Family Attached Unit (KTYPLVQT=3)	175	128	22	14
Building with 2-4 Units (KTYPLVQT=5)	499	256	67	29
Building with 5 or More Units (KTYPLVQT=6)	444	160	129	44

Table 4. Total Households in the Sample that Use Natural Gas, Have Good Data, Use Natural Gas for Main Heating Fuel, and Use Natural Gas for Main Heating Fuel and Have Good Data (Classified by Housing Type).

Housing Type	Total That Use Natural Gas	Total That Have Good Natural Gas Data	Total That Use Natural Gas for Main Heating Fuel	Total That Use Natural Gas for Main Heating Fuel and Have Good Natural Gas Data
Unknown	10	2	9	2
Mobile Home (KTYPLVQT=1)	81	44	74	42
Single-Family Detached Unit (KTYPLVQT=2)	1,641	1,285	1,510	1,188
Single-Family Attached Unit (KTYPLVQT=3)	141	92	127	82
Building with 2-4 Units (KTYPLVQT=5)	387	169	317	132
Building with 5 or More Units (KTYPLVQT=6)	317	69	201	47

Table 5. Total Households in the Sample that Use Fuel Oil, Have Good Data, Use Fuel Oil for Main Heating Fuel, and Use Fuel Oil for Main Heating Fuel and Have Good Data (Classified by Housing Type).

Housing Type	Total That Use Fuel Oil	Total That Have Good Fuel Oil Data	Total That Use Fuel Oil for Main Heating Fuel	Total That Use Fuel Oil for Main Heating Fuel and Have Good Fuel Oil Data
Unknown	-	-	-	-
Mobile Home (KTYPLVQT=1)	39	12	39	12
Single-Family Detached Unit (KTYPLVQT=2)	536	355	532	353
Single-Family Attached Unit (KTYPLVQT=3)	21	11	21	11
Building with 2-4 Units (KTYPLVQT=5)	99	25	97	25
Building with 5 or More Units (KTYPLVQT=6)	97	0	96	0

Table 6. Total Households in the Sample that Use Kerosene, Have Good Data, Use Kerosene for Main Heating Fuel, and Use Kerosene for Main Heating Fuel and Have Good Data (Classified by Housing Type).

Housing Type	Total That Use Kerosene	Total That Have Good Kerosene Data	Total That Use Kerosene for Main Heating Fuel	Total That Use Kerosene for Main Heating Fuel and Have Good Kerosene Data
Unknown	-	-	-	-
Mobile Home (KTYPLVQT=1)	33	21	32	21
Single-Family Detached Unit (KTYPLVQT=2)	53	32	49	32
Single-Family Attached Unit (KTYPLVQT=3)	-	-	-	-
Building with 2-4 Units (KTYPLVQT=5)	2	1	2	1
Building with 5 or More Units (KTYPLVQT=6)	2	1	1	1

Table 7. Total Households in the Sample that Use LPG, Have Good Data, Use LPG for Main Heating Fuel, and Use LPG for Main Heating Fuel and Have Good Data (Classified by Housing Type).

Housing Type	Total That Use LPG	Total That Have Good LPG Data	Total That Use LPG for Main Heating Fuel	Total That Use LPG for Main Heating Fuel and Have Good LPG Data
Unknown	-	-	-	-
Mobile Home (KTYPLVQT=1)	90	45	44	25
Single-Family Detached Unit (KTYPLVQT=2)	231	146	101	69
Single-Family Attached Unit (KTYPLVQT=3)	6	3	2	0
Building with 2-4 Units (KTYPLVQT=5)	16	4	11	2
Building with 5 or More Units (KTYPLVQT=6)	8	0	1	0

- We have good energy consumption data for most households living in single-family detached units but not for households living in multi-family units.

- In general, the LPG consumption data are not good.

- Most households that use natural gas use it as their main heating fuel. The same applies to fuel oil and kerosene.

Space-heating is an important part of any residential energy consumption model. In building regression models, it is wise to use only households for which we have good data. Therefore, the number of households in the sample that use a certain fuel for the main space-heating fuel and for which we have good data should be large if we are to construct a complete model for a particular fuel/housing type cell. The relatively large number of households of the above type for electricity/single-family detached units, natural gas/single-family detached units, and fuel oil/single-family detached units suggests that a reasonable model can be built for these cells from the NIECS data set. The models for electricity and natural gas were reviewed in Section 1. The fuel oil model is presented in the next section. For the other cells, it was decided to use a modification of one of the models built for the three cells above. If there were less than 10 households with good data in a cell, no model was constructed.

4. FUEL OIL MODEL FOR SINGLE-FAMILY DETACHED UNITS

The same technique was used to construct a regression model that expresses the predicted fuel oil consumption in MBtu for single-family detached units as was used to model the electricity and natural gas consumption. This included iterative fittings of least square solutions with all but the first solution using weighted least squares. The same outlier rejection technique was used. Four iterations were used in constructing the fuel oil model. Only households with good fuel oil data were used.

Originally, it was decided to use the same terms in the fuel oil consumption model as were used in the natural gas consumption model. This would facilitate comparisons between the two models.

Unfortunately, only two of the households with good fuel oil data did not use fuel oil as the main space-heating fuel. Both of these households used fuel oil as the fuel for a secondary space-heating system but not for water-heating. These two households would determine the coefficient for secondary space-heating and have no effect on any other coefficient in the model. Consequently, these households were dropped from the model building data set and the term for secondary space-heating was dropped from the model.

One of the households indicated that they used fuel oil for cooking. No other household used fuel oil for any appliance except water-heating. This single household was also dropped from the fuel oil model building data set. No term for fuel oil appliances other than water-heating were used in the model.

The term HFOMHT x NHEATDD x NROOMS was originally included in the model. The resulting coefficient for this term was .1865 with a standard deviation of .3042. Therefore, the coefficient is not significantly different from zero. The colinearity of the measures of size of the housing unit makes the estimates of the coefficients corresponding to the measures unstable. Consequently, the models, even if they have the same terms, will not be comparable term by term. For these reasons, the term HFOMHT x NHEATDD x NROOMS was dropped from the model.

The terms used in the final model for fuel oil consumption in single-family detached units are HFOWHT x NHSLDMEM, KMHTEQRC, HFOMHT x NHEATDD x NBATHRMS, HFOMHT x NHEATDD x NAGE01, HFOMHT x NHEATDD x KYHSBREC, HFOMHT x NHEATDD x NDRSAWS, and HFOMHT x NHEATDD x NSDRSAWS. Where HFOWHT is an indicator variable for water-heating with fuel oil and HFOMHT is an indicator variable for main space-heating with fuel oil. The rest of the variables are the same as those used in the natural gas consumption model. They are all listed in Appendix A.

Tables 8 and 9 summarize the results of the final fuel oil consumption model. Two households were rejected as outliers in fitting the model. These are listed in Appendix B. Figure 1 is a scatter plot of the residual for the

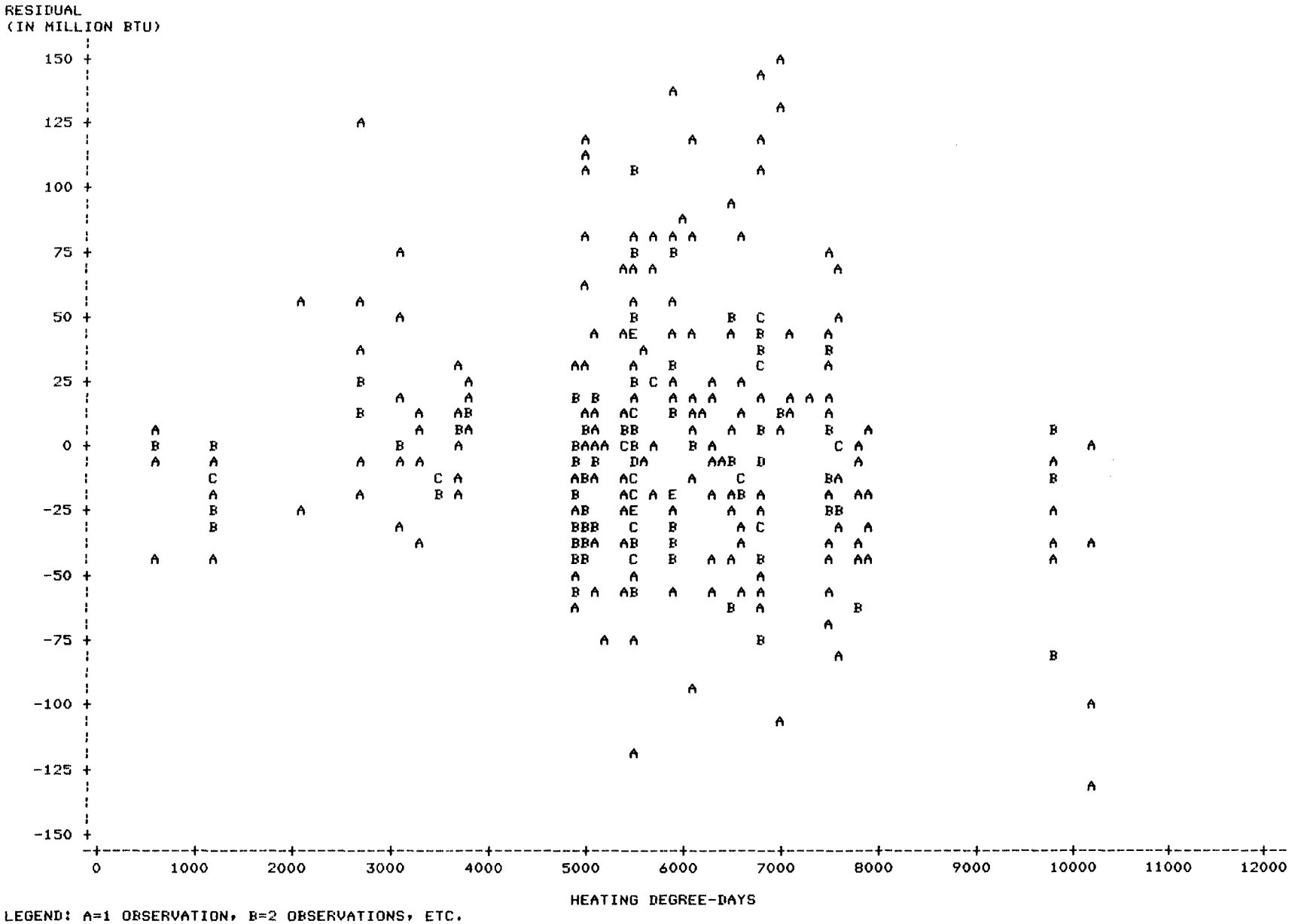
TABLE 8. ANALYSIS OF VARIANCE TABLE FOR FUEL OIL MODEL (USING NCFKYRB AS THE DEPENDENT VARIABLE).

SOURCE	DF	WEIGHTED SUM OF SQUARES	MEAN SQUARE	F VALUE	ATTAINED LEVEL OF SIGNIFICANCE	WEIGHTED R-SQUARE	WEIGHTED STANDARD DEVIATION
MODEL	8	7,822,798.49945054	977,849.81243132	67.06	0.0001	0.611390	120.75408
ERROR	341	4,972,307.97412489	14,581.54831122				
CORRECTED TOTAL	349	12,795,106.47357544					

TABLE 9. WEIGHTED LEAST SQUARE SOLUTION FOR FUEL OIL MODEL (USING NCFKYRB AS THE DEPENDENT VARIABLE).

PARAMETER	ESTIMATE	T-STATISTIC FOR H0: PARAMETER=0	ATTAINED 2-SIDED LEVEL OF SIGNIFICANCE	STD ERROR OF ESTIMATE
HFKWHT*NHSLDMEM	7,691.09303386	4.29	0.0001	1,792.53277393
KMHTEQRC				
RADIATOR OR HOT WATER PIPES (1)	80,537.59547152	9.17	0.0001	8,786.33497254
CENTRAL FORCED AIR (2)	39,711.26026684	6.67	0.0001	5,957.11506297
OTHER(3)	21,614.86923433	3.17	0.0017	6,815.56893710
HFKMHT*NHEATD*NBATHR	4.55597633	5.79	0.0001	0.78636840
HFKMHT*NHEATD*NAGE01	0.04407488	2.51	0.0126	0.01757194
HFKMHT*NHEATD*KYHSBR	-0.96883482	-4.36	0.0001	0.22235487
HFKMHT*NHEATD*NDRSAW	0.42827408	5.29	0.0001	0.08090706
HFKMHT*NHEATD*NSDRSA	-0.15121750	-2.27	0.0239	0.06667126

FIGURE 1. SCATTER PLOT OF RESIDUAL FOR FUEL OIL MODEL BY HEATING DEGREE DAYS.



LEGEND: A=1 OBSERVATION, B=2 OBSERVATIONS, ETC.

fuel oil model against NHEATDD and shows a drop at 9,000 heating degree-days. The same was shown in the corresponding plot for the natural gas model. Two logical explanations for this drop are better insulation and lower thermostat settings in the colder regions.

The value for mean square given in Table 8 can be used to give an approximate standard deviation to the estimated fuel oil consumption estimate. The approximate standard deviation equals the square root of (14,581.458 x YHAT), where YHAT is the estimated consumption. The form of the estimate comes from the use of 1/YHAT as the weight in the weighted least squares. For example, if YHAT equals 120,000, then an estimate of the standard deviation of YHAT is the square root of (120,000 x 14,581.548) which equals 41,830. This estimate is only approximate in that the error term is not normal, the weights are not exactly proportional to the variances, and the estimate does not take into account the clustering in the sample. In any case, it is evident that there is much variability in energy consumption among households with similar characteristics.

Table 10 displays the modeled fuel oil, natural gas, and electricity consumption for a hypothetical, but not necessarily typical, household. Each model has been estimated as if the particular fuel was used for water-heating and main space-heating. The natural gas and electricity models were adapted from Tables 7 and 8 of the report, National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, (DOE/EIA-0272). The appliance terms in the electricity consumption model were condensed into one term. The same was done for the air conditioning terms. The intercept term in the fuel oil model is contained in the term for the type of heating equipment. Some of the terms do not have values for all three models. In these cases, that particular term is not included in all of the consumption models.

Due to differences in the composition of the terms in the models and because of colinearity problems, the results are not comparable term by term. In fact, the totals are not strictly comparable. If the appliances and air conditioning terms are subtracted, then the adjusted totals should display the modeled water-heating and space-heating components. The resulting adjusted totals are 138,448 MBtu for fuel oil, 163,204 MBtu for natural gas, and 60,352 MBtu for electricity.

When comparing the adjusted totals, it is not clear that the intercept terms for the natural gas and electricity models belong in the water-heating plus space-heating components. For natural gas, the intercept term represents the baseline for all households. The same applies to electricity. Only households that heat with fuel oil were used in constructing the fuel oil model. Consequently, there is no intercept term for the fuel oil model.

The adjusted totals in themselves do not imply that it is more efficient to heat with electricity than fuel oil or, in turn, that it is more efficient than heating with natural gas. The Btu content of a fuel is measured as delivered to the household. But it takes several Btu of coal, oil, or natural

Table 10. Comparison of Estimated Fuel Oil, Natural Gas, and Electricity Consumption for a Hypothetical Household.

Independent Variable	Hypothetical Value	Fuel Oil Model	Contribution to Natural Gas Model (In MBtu)	Electricity Model
Intercept	-	0	7,886	1,839
Appliances	-	0	8,136	38,478
Hot Water Heating (NHSLDMEM x H__WHT)	4 x 1	30,764	32,733	14,271
Type of Heating Equipment (KMHTEQRC)	2	39,711	21,177	-
Electric Heat Pump (NHEATDD x HELMHT x HELHTPUM)	5,700 x 1 x 1	-	-	-14,646
Heating/Measure of Size				
NHEATDD x H__MHT x NBATHRMS	5,700 x 1 x 2	51,938	49,691	24,703
NHEATDD x H__MHT x NROOMS	5,700 x 1 x 7	-	45,810	-
NHEATDD x H__MHT x NDRSAWS	5,700 x 1 x 18	43,940	37,891	31,961
Age of Respondent/Heating				
NHEATDD x H__MHT x NAGE01	5,700 x 1 x 40	10,049	13,278	9,979
Code for Year House Built/Heating				
NHEATDD x H__MHT x KYHSBREC	5,700 x 1 x 5	-27,611	-32,114	-
Storm Doors and Windows/Heating				
NHEATDD x H__MHT x NSDRSAWS	5,700 x 1 x 12	-10,343	-13,148	-7,755
Air Conditioning	-	-	33,424	11,910
		138,448	204,764	110,740

gas to generate and deliver a Btu of electricity to a household. The difference between fuel oil and natural gas projected consumption statistics depends on the housing unit characteristics. If the heating systems for the hypothetical households were changed to systems that use radiators (KMHTEQRC = 1), the adjusted total for the fuel oil model would exceed the adjusted total for the natural gas model.

5. CONSUMPTION MODEL FOR OTHER HOUSING TYPES AND FUELS.

We decided that all housing type/fuel cells, except for single-family dwelling/electricity, natural gas, and fuel oil, did not have enough households with good data to adequately fit a model with similar variables as those used in fitting the three excepted models. For those cells, we decided to use the same technique, but to condense the variables into three or fewer terms. These terms would be linear combinations of the variables that are parts of the entire solution of the models in the large cells.

For the electricity model for households living in housing types other than single-family detached units, the model was based on the following three terms:

```
ELBASE =  1,838.90
                  +1,542.91 x NELCKDV
                   +249.498 x KINCOME
                   -46.1915 x NAGE01
                  +1,800.27 x NHSLDMEM
                  +3,567.84 x HELWHT x NHSLDMEM
                  +5,010.04 x HELWHT x HELDISHW
                  +3,176.97 x NELFRIG
                  +3,246.81 x HELDISHW
                  +3,827.05 x HELCLSDY
                  +4,456.46 x HSPFDFRZ
                  +2,441.15 x HAUTOWSH
                   +108.336 x NTOTAL
                  +3,282.23 x HSBNMELH

ELHEATDD =     -2.56940 x NHEATDD x HELMHT x HELHTPUM
              +0.0437668 x NHEATDD x HELMHT x NAGE01
                +2.16693 x NHEATDD x HELMHT x NBATHRMS
               -0.113379 x NHEATDD x HELMHT x NSDRSAWS
               +0.311507 x NHEATDD x HELMHT x NDRSAWS

ELCOOLDD =      0.697939 x NCOOLDD x NRMELRAC
              +0.0702737 x NCOOLDD x NRMELRAC x KINCOME
               +0.412585 x NCOOLDD x NRMELCAC
               +0.125661 x NCOOLDD x NRMELCAC x KINCOME
```

Note that these terms were derived intact from the solution to the electricity consumption model for single-family detached units. In using these three terms as the independent variables, we are assuming that ELBASE represents an index for electricity end uses that are not related to heating or cooling

degree-days, ELHEATDD is an index for electricity demand for primary space-heating and ELCOOLDD is an index for electricity demand for cooling. The only other change in the model building technique was that the intercept was arbitrarily set equal to zero.

For the other fuel types, the heating component of the consumption models overshadows the other parts. For this reason, only one term was used in the cells with small numbers of households with good data. For natural gas and LPG, the single term used was the projected natural gas consumption given by the model for single-family detached dwellings. This projection is given by:

```
NGMODEL =   7,885.86
                    +478.570 x NGASNDX
                  +8,183.13 x NHSLDMEM x HGASWHT
                  +39,380.1 x HSBNMGSH
                  +35,976.6 x HGSHTEQ1
                  +21,177.3 x HGSHTEQ2
                  +7,946.84 x HGSHTEQ3
                    +4.3589 x HGASMHT x NHEATDD x NBATHRMS
                    +0.0582 x HGASMHT x NHEATDD x NAGE01
                    -1.1268 x HGASMHT x NHEATDD x KYHSBREC
                    +1.1481 x HGASMHT x NHEATDD x NROOMS
                    +0.3693 x HGASMHT x NHEATDD x NDRSAWS
                    -0.1922 x HGASMHT x NHEATDD x NSDRSAWS
                    +3.9791 x NCOOLDD x NRMGASAC.
```

Here, HGSHTEQ1 is an indicator variable for gas main heating using radiators or hot water pipes; HGSHTEQ2 is an indicator variable for gas main heating using central forced-air; and HGSHTEQ3 is an indicator variable for other types of gas main heating equipment. For fuel oil and kerosene, the single term used was the projected fuel oil consumption given by the model for single-family detached dwellings. This projection is given by:

```
FOMODEL =  7,691.09 x HFOWHT x NHSLDMEM
          +80,537.60 x HFOMHT x HHEATEQ1
          +39,711.26 x HFOMHT x HHEATEQ2
          +21,614.87 x HFOMHT x HHEATEQ3
             +4.5560 x HFOMHT x NHEATDD x NBATHRMS
             +0.0441 x HFOMHT x NHEATDD x NAGE01
             -0.9688 x HFOMHT x NHEATDD x KYHSBREC
             +0.4283 x HFOMHT x NHEATDD x NDRSAWS
             -0.1512 x HFOMHT x NHEATDD x NSDRSAWS.
```

Here, HHEATEQ1 is an indicator variable for main space-heating using radiators or hot water pipes; HHEATEQ2 is an indicator variable for main space-heating using central forced-air; and HHEATEQ3 is an indicator variable for other types of main space-heating equipment. For kerosene, the terms HFOWHT and HFOMHT were changed to indicators for kerosene usage. In these cases, NGMODEL and FOMODEL represent indexes for potential energy use.

The same technique was used as before including setting the intercept equal to zero.

Table 11 summarizes the results for the additional cells. The coefficients and their standard errors were obtained from the results of the final weighted least square regressions of the modeling procedures. A list of households rejected as outliers is given in Appendix B.

The standard errors shown in Table 11 are probably conservative in that they do not take the clustering in the sampling plan into account. They are shown in the table to give a bottom line to the precision of the estimates.

The listed coefficients in Table 11 are estimates of the proportion of the single-family modeled consumption that we should expect in other cells. For example, if the single-family electricity consumption model gives values of ELBASE = 30,000 MBtu, ELHEATDD = 50,000 MBtu and ELCOOLDD = 20,000 MBtu, then an estimate of the electricity consumption for a mobile home with the same household and housing characteristics would be (30,000)(.96) + (50,000)(1.00) + (20,000)(1.32) = 105,200 MBtu. If the single-family natural gas consumption model gives an estimated consumption of 125,000 MBtu, then an estimate for the LPG consumption for a household living in a single-family detached dwelling that uses LPG for all of the same end uses and has the same household and housing characteristics would be (125,000)(.81) = 101,250 MBtu.

The standard deviation of an estimated consumption for a household whose characteristics place it in a small cell should reflect not only the standard deviation of the proportion but also the standard deviation of the corresponding large cell consumption estimate. The impact of the second standard deviation tends to overshadow the first one.

The following general trends are illustrated in Table 11.

- Households living in large apartment buildings and paying their own utility bills tend to use less energy than households living in other housing types whose household characteristics are the same.

- The cooling degree-day component of the modeled yearly electricity consumption behaves differently than the heating degree-day component and the baseline component. In particular the coefficient for the cooling degree-day component tends to be larger than the other two components and larger than 1.0.

- Households that use LPG tend to use less energy than households with the same characteristics that use natural gas.

Table 11. Estimates of Coefficients (with Standard Errors).

Main Heating Fuel	Housing Type	ELBASE	ELHEATDD	ELCOOLDD	NGMODEL	FQMODEL
Electricity	1 (Mobile Home)	.96 (.04)	1.00 (.06)	1.32 (.20)		
	3 (Single-Family Attached)	.84 (.03)	1.35 (.11)	1.31 (.16)		
	5 (Small Apartment Building)	.86 (.03)	.80 (.09)	1.15 (.17)		
	6 (Large Apartment Building)	.68 (.03)	.71 (.05)	1.17 (.17)		
Natural Gas	1				1.02 (.05)	
	3				.90 (.04)	
	5				1.02 (.03)	
	6				.59 (.03)	
Fuel Oil	1					1.01 (.11)
	3					1.01 (.11)
	5					1.03 (.08)
LPG	1				.77 (.06)	
	2				.81 (.03)	
Kerosene	1					1.01 (.06)
	2					1.10 (.13)

6. DISCUSSION OF RESULTS

The reader must be careful not to place too much emphasis on the individual terms in the models. The coefficients for individual terms may be drastically altered by changing the set of independent variables in the regression equation or transforming the dependent variable. As a consequence, the coefficient for a particular independent variable could easily include the effect of other variables that are correlated with it. This applies to the three basic models as well as the modified models. It is much more relevant to consider the modeled total annual consumption of a hypothetical household as a value in itself and not as the sum of well-defined individual terms. The total itself is only meaningful for households whose independent variables are within the range used in constructing the model. The modeled consumption for a hypothetical two-person household living in a house with eight bathrooms and 22 other rooms may not be relevant.

The problems in using the individual terms in the consumption models can be illustrated by examining the term for the number of electric refrigerators in the electricity consumption model. The households used in constructing the electricity consumption model for households living in single-family detached houses can be broken down as follows: five of the households had no refrigerators; 1,735 had one; 393 had two; and 14 had three. Therefore, the coefficient for the term NELFRIG is based mainly on the comparison between the 1,735 households with one refrigerator and the 393 households with two. It might be argued that the term applies mainly to the second refrigerator. The first refrigerator in the data set is the one the household designated as the most-used. In addition, the coefficient for NELFRIG may be measuring some other difference between the energy consumption of one- and two-refrigerator households. Even if the coefficient measured only the electricity consumption due to refrigerators, it would only be an average consumption. Refrigerators vary in size, age, design, and energy consumption. The model uses one value for all refrigerators.

The coefficients obtained for the modified models apply only to the linear combinations and not to the individual terms in the linear combinations. For example, the coefficients for NGMODEL in the natural gas consumption model for households living in large apartment buildings is .59. This does not imply that households living in apartments use 41 percent less hot water per capita than households living in single-family detached houses. It only implies that households paying their own utility bills and living in large apartment buildings use approximately 41 percent less natural gas annually than comparable households living in single-family detached homes.

Appendix A

NIECS PUBLIC USE FILE VARIABLES USED IN CONSUMPTION MODELS

Table A1 lists the household variables used in the construction of the consumption models. The variable names and values are documented in the NIECS public use file. Table A2 defines the variables used in the modeling process that are transformations of the variables in Table A1. Table A3 gives the codes for the variable KINCOME, household income, and the variable KYHSBREC, year house built. Table A4 gives the gas appliance index values for individual appliances. The value of NGASNDX, the gas appliance index, is the sum of the index values over all gas appliances in the housing unit.

Table A1. **Household** Variables Selected for Use in the Analysis of Natural **Gas, Electricity**, Fuel Oil, Kerosene, and LPG Consumption.

Variable Name	Variable Description
HAUTOWSH	Indicator Variable for Automatic Washing Machine
HCENTAC	Indicator Variable for Central Air Conditioning
HELCLSDY	Indicator Variable for Electric Clothes Dryer
HELDISHW	Indicator Variable for Electric Dish Washer
HELHTPUM	Indicator Variable for Electric Heat Pump
HELOVEN	Indicator Variable for Electric Oven
HELRANGE	Indicator Variable for Electric Range
HROOMAC	Indicator Variable for Electric Room Air Conditioners
HSPFDFRZ	Indicator Variable for Separate Food Freezer
KFLCNAC	Code for Central Air Conditioning Fuel
KFLMHEAT	Code for Main Heating System Fuel
KFLSHEAT	Code for Secondary Heating System Fuel
KINCOME	Code for Household Income Level (see Table A3)
KMHEATEQ	Code for Main Heating Equipment
KREFRIG1	Code for First Refrigerator Fuel
KREFRIG2	Code for Second Refrigerator Fuel
KWHEATFL	Code for Water Heating Fuel
KYHSBREC	Code for Year House Built (see Table A3)
NAGE01	Age of Respondent
NCELYRB	Annual Consumption of Electricity in Thousands of Btu

Table A1. Household Variables Selected for Use in the Analysis of Natural Gas, Electricity, Fuel Oil, Kerosene, and LPG Consumption. (Continued)

Variable Name	Variable Description
NCFKYRB	Annual Consumption of Fuel Oil and/or Kerosene in Thousands of Btu
NCLPYRB	Annual Consumption of LPG in Thousands of Btu
NCNGYRB	Annual Consumption of Natural Gas in Thousands of Btu
NCOMBATH	Number of Complete Bathrooms
NCOOLDD	Number of Cooling Degree-Days Based on 40-Year Average
NDOORS1	Number of Outside Doors and Sliding Glass Doors
NGASNDX	Gas Appliance Index (see Table A4)
NHAFBATH	Number of Half Bathrooms
NHEATDD	Number of Heating Degree-Days Based on 40-Year Average
NHSLDMEM	Number of Household Members
NREFRIG	Number of Refrigerators
NROOMAC	Number of Rooms Air Conditioned
NROOMS	Number of Rooms in House
NSDOORS	Number of Storm Doors
NSTRMWIN	Number of Storm Windows
NSWINSGD	Number of Storm Sliding Glass Doors
NUMWINDS	Number of Windows

Table A2. Household Variables Constructed For Use in the Analysis of Natural Gas, Electricity, Fuel Oil, Kerosene, and LPG Consumption.

Variable	Definition		Description
HELFRIG1	1	if KREFRFL1 = 1	Indicator variable for electric refrigerator.
	0	otherwise	
HELFRIG2	1	if KREFRFL2 = 1	Indicator variable for electric refrigerator
	0	otherwise	
HELFRIG3	1	if NREFRIG = 3 and KREFRFL1 = 1	Indicator variable for electric refrigerator.
	0	otherwise	
HELHTPUM	1	if KMHEATEQ = 4	Indicator variable for electric heat pump.
	0	otherwise	
HELMHT	1	if KFLMHEAT = 5	Indicator variable for electric main space-heating.
	0	otherwise	
HELWHT	1	if KWHEATFL = 5	Indicator variable for electric water heating.
	0	otherwise	
HFOMHT	1	if KFLMHEAT = 3	Indicator variable for fuel oil main space-heating
	0	otherwise	
HFOWHT	1	if KWHEATFL = 3	Indicator variable for fuel oil water heating.
	0	otherwise	
HGASMHT	1	if KFLMHEAT = 1	Indicator variable for natural gas main space-heating.
	0	otherwise	
HGASWHT	1	if KWHEATFL = 1	Indicator variable for natural gas water heating.
	0	otherwise	

Table A2. Household Variables Constructed For Use in the Analysis of Natural Gas, Electricity, Fuel Oil, Kerosene, and LPG Consumption (Continued).

Variable	Definition		Description
HKERMHT	1 0	If KFLMHEAT = 4 otherwise	Indicator variable for kerosene main space-heating.
HKERWHT	1 0	if KWHEATFL = 4 otherwise	Indicator variable for kerosene water heating.
HLPMHT	1 0	If KFLMHEAT = 2 otherwise	Indicator variable for LPG main space-heating.
HLPWHT	1 0	if KWHEAT = 2 otherwise	Indicator variable for LPG water heating.
HSBNMELH	1 0	if KFLSHEAT = 5 and KFLMHEAT ≠ 5 otherwise	Indicator variable for electric secondary but not main space-heating.
HSBNMGSH	1 0	if KFLSHEAT = 1 and KFLMHEAT ≠ 1 otherwise	Indicator variable natural gas secondary but not main space-heating.
HSBNMLPH	1 0	if KFLSHEAT = 2 and KFLMHEAT ≠ 2 otherwise	Indicator variable for LPG secondary but not main space-heating.
KMHTEQRC	1 2 3 4	if KFLMHEAT = 1 and KMHEATEQ = 1 or 2 if KFLMHEAT = 1 and KMHEATEQ = 3 if KFLMHEAT = 1 and KMHEATEQ ≠ 1, 2, and 3 if KFLMHEAT ≠ 1	Class variable for gas main heating equipment. 1 - System where heat is distributed by water 2 - Central forced-air 3 - Other (usually space-heater) 4 - Nongas main heating

Table A2. Household Variables Constructed For Use in the Analysis of Natural Gas, Electricity, Fuel Oil, Kerosene, and LPG Consumption (Continued).

Variable	Definition	Description
NBATHRMS	NCOMBATH + 1/2 (NHAFBATH)	Number of bathrooms.
NDRSAWS	NDOORS1 + NUMWINDS	Number of doors and windows.
NELCKDV	HELOVEN + HELRANGE	Number of major electric cooking appliances.
NELFRIG	HELFRIG 1 + HELFRIG 2 + HELFRIG 3	Number of electric refrigerators.
NRMELCAC	0 if HCENTAC = 0 0 if HCENTAC = 1 and KFLCNAC = 1 NROOMAC-HROOMAC if HCENTAC = 1 and KFLCNAC = 2	Number of rooms potentially air-conditioned by electric central units.
NRMELRAC	0 if HROOMAC = 0 1 if HROOMAC = 1 and HCENTAC = 1 NROOMAC if HROOMAC = 1 and HCENTAC = 0	Number of rooms potentially air-conditioned by electric room units.
NRMGASAC	0 if HCENTAC = 0 0 if HCENTAC = 1 and KFLCNAC = 2 NROOMAC-HROOMAC if HCENTAC = 1 and KFLCNAC = 1	Number of rooms potentially air-conditioned by natural gas central units.
NSDRSAWS	NSDOORS + NSWINSGD + NSTRMWIN	Number of storm doors and windows.
NTOTAL	NROOMS + NDOORS1 + NUMWINDS	Overall measure of the size of the house.

Table A3. Code for Household Income and Year House Built.

Variable	Value	Code
KINCOME	1	Under $3,000
(Household Income)	2	$3,000 - $4,999
	3	$5,000 - $7,999
	4	$8,000 - $9,999
	5	$10,000 - $11,999
	6	$12,000 - $14,999
	7	$15,000 - $19,999
	8	$20,000 - $24,999
	9	$25,000 - $29,999
	10	$30,000 - $34,999
	11	$35,000 - $39,999
	12	$40,000 - $44,999
	13	$45,000 - $49,999
	14	$50,000 or more
KYHSBREC	1	In 1939 or earlier
(Year House Built)	2	In 1940 - 1949
	3	In 1950 - 1959
	4	In 1960 - 1964
	5	In 1965 - 1969
	6	In 1970 - 1974
	7	In 1975 or later

Table A4. Gas Appliance Index.

Appliance	Index Value
Gas Clothes Dryer	7
Gas Oven/Range	10
Gas Refrigerator	12
Outdoor Gas Light	18

Appendix B

OUTLIERS DETECTED IN ESTIMATING CONSUMPTION MODELS

The outliers detected in constructing the electricity and natural gas consumption models for households living in single-family detached houses are listed in Tables 3 and 6 of National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, (DOE/EIA-0272). The outliers detected in constructing the other models are listed in Table B1.

Table B1. Households that Were Detected as Outliers.

Fuel	Housing Type	Household ID of Outliers (NSQIDDOE)
Fuel Oil	Single-Family Detached	1,278
		3,772
Electricity	Mobile Home	2,918
		3,155
		3,248
		3,591
		3,799
		3,982
		4,829
Electricity	Single-Family Attached	1,174
		2,648
		2,854
		3,757
		4,688
Electricity	Buildings with 2-4 units	1,359
		1,597
		2,123
		3,897
		3,946
Electricity	Buildings with 5 or more units	3,102
		3,310
LPG	Single-Family Detached	2,520

*U.S. GOVERNMENT PRINTING OFFICE : 1981 0-361-068/1201

Other Reports Published by the Office of Energy Markets and End Use, End Use Energy Division:

Residential Energy Consumption Survey: Conservation, Feb. 1980, DOE/EIA-0207/3, GPO Stock No. 061-003-00087-8, $6.00.

Single-Family Households: Fuel Inventories and Expenditures: National Interim Energy Consumption Survey, Dec. 1979, DOE/EIA-0207/1, GPO Stock No. 061-003-00075-4, $1.75.

Residential Energy Consumption Survey: Characteristics of the Housing Stocks and Households, 1978, Feb. 1980, DOE/EIA-0207/2, GPO Stock No. 061-003-00093-2, $4.25.

Residential Energy Consumption Survey: Consumption and Expenditures, April 1979 Through March 1979, July 1980, DOE/EIA-0207/5, GPO Stock No. 061-003-00131-9, $6.50.

Residential Energy Consumption Survey: Consumption Patterns of Household Vehicles, June to August, 1979, June 1980, DOE/EIA-0207/4, GPO Stock No. 061-003-00156-4, $3.75.

Nonresidential Buildings Energy Consumption Survey: Building Characteristics, March 1981, DOE/EIA-0246, GPO Stock No. 061-003-00171-8, $5.50.

Nonresidential Buildings Energy Consumption Survey: Fuel Characteristics and Conservation Practices, June 1981, DOE/EIA-0278, GPO Stock No. 061-003-00200-5, $8.00.

Residential Energy Consumption Survey: 1979-1980 Consumption and Expenditures, Part I: National Data (Including Conservation), April 1981, DOE/EIA-0262/1, GPO Stock No. 061-003-00191-2, $5.50.

Residential Energy Consumption Survey: 1978-1980 Consumption and Expenditures, Part II: Regional Data, May 1981, DOE/EIA-0262/2, GPO Stock No. 061-003-00189-1, $8.50.

The National Interim Energy Consumption Survey: Exploring the Variability in Energy Consumption, July 1981, DOE/EIA-0272, GPO Stock No. 061-003-00205-6, $4.25.

Copies of the above reports are available from:

Superintendent of Documents
U.S. Government Printing Office
Washington, D.C. 20402

National Interim Energy Consumption Survey: Household Interview File, DOE/DF-81-001 (magnetic tape), available from the National Technical Information Service - Computer Products Division, 5285 Port Royal Road, Springfield, Virginia, Accession No. PB-81-108714, $125.00.

U.S. Department of Energy
Energy Information Administration
National Energy Information Center, EI-20
Forrestal Building
Washington, D.C. 20585