Title

A Review of Assumptions and Analysis in EPRI EA-3409, "Household Appliance Choice: Revision of REEPS Behavioral Models"

Permalink

https://escholarship.org/uc/item/5v24b6hr

Authors

Wood, D J
Ruderman, H
McMahon, J E

Publication Date

1989-05-01

Copyright Information

LBL-20332 c.2
UC-350

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

APPLIED SCIENCE DIVISION

A Review of Assumptions and Analysis in EPRI EA-3409, "Household Appliance Choice: Revision of REEPS Behavioral Models"

D.J. Wood, H. Ruderman, and J.E. McMahon

May 1989

APPLIED SCIENCE DIVISION

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

DISCLAIMER

LBL-20332

A REVIEW OF ASSUMPTIONS AND ANALYSIS IN EPRI EA-3409: "HOUSEHOLD APPLIANCE CHOICE: REVISION OF REEPS BEHAVIORAL MODELS"*

David J. Wood, Henry Ruderman and James E. McMahon

Energy Analysis Program
Applied Science Division
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

May 1989

ABSTRACT

This paper revises and extends EPRI report EA-3409, "Household Appliance Choice: Revision of REEPS Behavioral Models." That paper reported the results of an econometric study of major appliance choice in new residential construction. Errors appeared in two tables of that report. We offer revised versions of those tables, and a brief analysis of the consequences and significance of the errors.

The present paper also proposes several possible extensions and re-specifications of the models examined by EPRI. Some of these are judged to be highly successful; they both satisfy economic intuition more completely than the original specification and produce a better quality fit to the dependent variable. We feel that inclusion of these modifications produces a more useful set of coefficients for economic modeling than the original specification.

* This work was supported by the Assistant Secretary for Conservation and Renewable Energy, Office of Building and Community Systems, Building Equipment Division of the U.S. Department of Energy, under Contract No. DE-AC03-76SF00098. The authors also wish to thank the Electric Power Research Institute for permission to quote extensively from their report EPRI EA-3409.

A REVIEW OF ASSUMPTIONS AND ANALYSIS IN EPRI EA-3409: "HOUSEHOLD APPLIANCE CHOICE: REVISION OF REEPS BEHAVIORAL MODELS"

David J. Wood, Henry Ruderman and James E. McMahon

1. Introduction

The Electric Power Research Institute (EPRI) directs the operation of an energy demand forecasting model, the Residential End-use Energy Planning System (REEPS). EPRI commissioned Cambridge Systematics, Inc. to estimate from national survey data consumer appliance choice as a function of appliance and household characteristics. The final report of that effort appeared as EPRI EA-3409 in February 1984, titled "Household Appliance Choice: Revision of REEPS Behavioral Models."

This paper focuses on EPRI's models of residential space heating technology choice. That choice was modeled as a nested logit structure, with consumers choosing whether to have central air conditioning or not, and, given that choice, what kind of space heating system to have. The model included five space heating alternatives with central cooling (gas, oil, and electric forced-air; heat pumps; and electric baseboard) and eight alternatives without it (gas, oil, and electric forced-air; gas and oil boilers and non-central systems; and electric baseboard heat). The structure of the nested logit model is shown in Figure 1 below.

Two of the tables appearing in EPRI's report (Tables 8 and 10) contained errors. The problems in Table 8 were negligible, resulting from the inclusion in the dataset of some households with erroneous or missing data. But the errors in Table 10 were more serious, with consequences to the elasticities estimated from the model. Problems in Table 10 resulted from the accidental miscalculation of one of the independent variables in the model.

This paper also considers several extensions and modifications of the logit models estimated by EPRI. Those modifications are considered separately below.

Heat Pump Cost Calculations

The EPRI models included capital and operating costs of each alternative technology as independent variables. These costs were calculated using a model of residential thermal integrity and heating loads to estimate the required capacity and expected fuel consumption of each technology in each household. Local construction costs and fuel prices were used to translate these numbers into dollar values.

This process seems to have incurred some errors in the calculation of heat pump capital and operating costs. In particular, heat pump capital costs seem quite high (on average, about four times the cost of a conventional air conditioner for the same household), and operating costs somewhat low. Both of these errors could be explained by an oversizing of heat pumps in relatively cold climates. Actual construction practice seldom calls for installing a heat pump in locations with a severe winter climate, but if it were so installed, standard practice is to size it for the summer cooling load and supplement the winter heating load with electric resistance heat (essentially, an electric forced-air heating system).

To rectify the apparent overstatement of capital costs, we propose a conservative upper bound for any household's heat pump capital costs, and replace the original heat pump capital costs with the bound for all households violating it. We also replace the operating cost of heat pumps for those households with a "corrected" calculation. We then re-estimated the two logit regressions affected by the change, and found significant improvements in the overall quality of the fit of the model to the data.

Cumulative Gas Restrictions

The period covered by the EPRI study included several major disruptions in the U.S. energy supply. In particular, many gas utilities were obliged to restrict or even completely prohibit new residential gas service in that period. These restrictions had a significant effect on the market for space heating technologies. EPRI modeled these effects by including binary variables which indicated the presence or absence of three different levels of restrictions for the gas utility serving each household in the year the house was built. The effects of these variables were all found to be statistically significant.

We wished to investigate whether the effects of gas restrictions extended beyond the period during which they were in force. This could occur either because of a psychological mechanism, as consumers and builders remember the prior curtailments and wish to avoid that possibility in the future, or because of the growth of a sales and service infrastructure for alternative technologies during the curtailment.

We tested this idea by extending the definition of the "gas restrictions" variables so that they were no longer binary (absent = 0, present = 1), but could also be fractional if restrictions had been present in years prior to the construction of a particular residence. We took into account both the duration of the restriction and the time since it was last in effect. Changing the gas restrictions variables in this way significantly improved the quality of the fit of the model to the data.

Fuel Price Expectations

EPRI's models of space heating choice included operating costs as independent variables. These costs were based on the actual fuel prices faced by the consumer in the year the residence was built. What if consumers based their choice, not on the actual prices, but on some form of anticipated prices?

We tested this by noting the one-year change in each fuel price for each consumer and then using this in several forms to calculate one-year, five-year, and ten-year anticipated prices. In all cases, the models based on anticipated prices were inferior to the equivalent specification with actual prices. Although the idea that consumers have price expectations and use them in making economic choices continues to be attractive, it is not supported by the relatively simple model considered here.

Discount Rates Varying with Household Characteristics

In models using a linear combination of exogenous variables to estimate consumer utility, the ratio of the capital cost coefficient to operating cost coefficient gives an implicit discount rate for the choices reported in that data set. Standard economic theory suggests that this discount rate should decline with increasing income, and this result may be observed by including the product of operating cost and income as an exogenous variable of the model. EPRI did this in both of the two heating choice sub-models, finding the expected result in one case and a counterintuitive result (discount rates rising with increasing income) in the other. In that case, the model was rejected in favor of using only operating cost (and not operating cost times income) as an exogenous variable.

We chose to consider other reasonable characteristics on which a household's discount rate might vary. One extension was to include the product of operating cost and maximum heat load as an independent variable. Heat load is positively correlated with income (higher income households tend to be larger, which leads to higher maximum heat loads), and both capital and operating costs (larger houses again). We found that, for the sub-model of heating choice given central cooling, formulations which included operating cost and heat load consistently out-performed (in the statistical sense) formulations using operating cost and income or operating cost alone. Essentially, the data indicate that households with large heating needs (either because of their size, or their geographic location) tend to pay more attention to considerations of future operating cost than do households with small heating loads, irregardless of income.

Another extension was to consider the product of operating cost and EPRI's measure of summer climate (which captures a sense of air conditioning desirability) in the central cooling choice model. Alternative formulations of the cooling choice model which used that product of variables consistently out-performed (again, in the statistical sense) those which included only a pure operating cost effect, regardless of the choice of variables in the sub-models of heating choice.

The present paper is organized as follows: Section 2 discusses the errors in EPRI's Tables 8 and 10. Tests of the significance of these differences relative to their uncertainties are also presented. Section 3 examines the consequences of these errors to elasticities and predicted market shares estimated from the model. Section 4 reports in detail the extensions and modifications to EPRI's models discussed immediately above. A brief set of conclusions and references to papers mentioned in the text follow.

There is a single appendix, which explains the statistical test used to determine the significance of differences between alternative sets of coefficients.

2. Errors in EPRI's Tables 8 and 10

Errors in EPRI's Table 8

The problem in EPRI's Table 8 is that six of the 842 households in that regression have an "income" value of zero. Even if this represents a true value of no income, it is doubtful that this zero-income status is relevant in the selection of space heating technology. It is more likely either a temporary condition, or an erroneous value representing missing data. In either case those households should be excluded from the regression. The numbers reported in EPRI's Table 8 are based on including the zero-income households.[1] Excluding them produces slightly different coefficient estimates, as shown in the table below. Evidence that EPRI considered the revised formulation (zero-income households excluded) correct is that coefficients from it were used to calculate the "inclusive value" term for space heat given central cooling.

We can test the significance of the difference of the two vectors of estimated coefficients using the standard likelihood ratio test, where the null hypothesis is that the true value of the revised parameters, calculated on the revised data set, are equal to the maximum likelihood estimate of the original parameters, which are treated as constants. This test is discussed more fully in Appendix 1.

Table I
Comparison of Original and Revised Versions of EPRI's Table 8

Effect of Zero-Income Households on Choice of Space Heating (given central cooling)

	Original: zero-income households included		Revised: zero-income households excluded	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-0.1645	(-6.646)	-0.1657	(-6.640)
normalized operating cost	-2.515	(-4.884)	-2.473	(-4.757)
operating cost × income	-0.0438	(-2.955)	-0.04503	(-3.009)
gas restrictions type 1	-2.352	(-7.568)	-2.428	(-7.598)
gas restrictions type 2	-0.9997	(-2.206)	-1.001	(-2.209)
gas restrictions type 3	-1.801	(-6.042)	-1.793	(-5.999)
heat pump trend	0.04591	(3.943)	0.04506	(3.855)
oil forced-air choice	-1.827	(-8.756)	-1.827	(-8.750)
elec forced-air choice	1.007	(3.882)	0.9907	(3.811)
heat pump choice	-0.4085	(-1.765)	-0.3947	(-1.700)
elec baseboard choice	-1.098	(-4.170)	-1.097	(-4.170)
Likelihood Ratio Test				
log likelihood:	-938.0		-932.4	
log likelihood revised model restricted to original results:			-932.5	
2 × difference ($\approx \chi^2_{11}$)			0.2	

[1] There was also a minor typographical error in EPRI's reporting of the "zero-income households included" version of Table 8. That error is corrected in this paper.

Working with the revised data set (zero-income households excluded), we can calculate the log likelihood of each set of parameters. Twice the difference in log likelihood of the two parameter sets is called the "likelihood ratio statistic" and, under a null hypothesis of no difference between the two vectors of parameter estimates, is asymptotically distributed as a χ^2 with degrees of freedom equal to the number of parameters estimated (eleven here). The difference in the log likelihood of these two parameter sets is only 0.1, calculated on the revised data set. A χ^2_{11} statistic of 0.2 (twice 0.1) is not significant, so we must conclude that the change produced by excluding the zero-income households is insignificant compared to the uncertainty in the estimated parameters.

Most of the increase in log-likelihood (from -938.0 to -932.4) that occurs when we exclude zero-income households is due to the fact that the calculation is carried out over fewer households, leaving less total contribution to total log-likelihood than originally. Thus, the revised model appears to be a non-trivial improvement (gain of 5.6 in the log-likelihood), when in fact, the gain is mostly illusory. The comparison of original and revised coefficients on the reduced (i.e., revised) dataset reveals the true degree of difference between the two sets of parameters.

Errors in EPRI's Table 10

The problems in Table 10 need to be understood with some background in the nested logit calculation. The following passage from EPRI's report[2] describes the model structure:

> Space heating and air conditioning decisions are modeled jointly with the dependent variable representing the choice of a space heat/air conditioning combination. The empirical specification is a generalized version of the multinomial logit known as the nested logit model. This functional form allows differential substitution between alternative appliance combinations. The multinomial logit does not allow such differential substitution since the relative odds of any two alternatives are independent of the availability and attributes of other alternatives. This is implausible for some applications. Consider, for example, an air conditioning market in which a consumer could choose between a conventional unit or no unit at all. Suppose a new technology--heat pumps--is introduced to the market. According to the multinomial logit model, the new technology would draw its market share proportionately from non-buyers and buyers of conventional units. Common sense suggests that the heat pump alternative would draw its market share primarily from conventional air conditioning because of their similarity.
>
> In order to capture this differential substitution, the nested logit model groups similar alternatives in a hierarchical structure. In the example of central air conditioning, the multinomial logit representation would treat the three alternatives (no a/c, conventional a/c, and heat pump) equivalently.... [In the nested logit representation,] the two central cooling options are grouped on a separate branch of the probability tree reflecting their close substitutability. Conceptually, the choice is broken down into two levels. At the upper level, the choice alternatives are cooling and no cooling. At the lower level, the alternatives are conventional air conditioning and heat pump, **given** the cooling choice. The decision at each level is represented by a multinomial logit model. The values of the explanatory variables in the lower level choice depend upon the branch of the probability tree. At the upper level, an index of the aggregate characteristics of the lower level alternatives is used as an explanatory variable in the choice model.
>
> For the application of the nested logit to the joint space heating and air conditioning decision, we specify the following general form for utility:
>
> $$U_{ij} = V_i + W_{ij} + \epsilon_{ij}$$
>
> where i indexes central air conditioning alternatives;
> j indexes space heating alternatives;
> U_{ij} is the total utility from a given ij combination;

[2] Electric Power Research Institute [1984] page 3-3

V_i is the the typical or representative utility from central air conditioning alternative i;

W_{ij} is the the representative utility from air conditioning and space heating alternative ij;

ϵ_{ij} are random components of utility reflecting unobserved characteristics and random consumer tastes.

Following McFadden [1978], if the random terms are assumed to be distributed according to the following form of the Generalized Extreme Value (GEV) distribution:

$$F(\epsilon_{11}, \ldots, \epsilon_{ij}) = \exp\left\{ - \sum_i \left[\sum_j e^{\epsilon_{ij}/(1-\theta)} \right]^{(1-\theta)} \right\}$$

Then the conditional and marginal probabilities $P_{j|i}$ and P_i can be expressed in the following closed forms:

$$P_{j|i} = \frac{e^{W_{ij}/(1-\theta)}}{\sum_{j'} e^{W_{ij'}/(1-\theta)}}$$

and

$$P_i = \frac{e^{V_i+(1-\theta)J_i}}{\sum_{i'} e^{V_{i'}+(1-\theta)J_{i'}}}$$

where

$$0 \leq \theta \leq 1$$

and

$$J_i = \ln\left(\sum_{j'} e^{W_{ij'}/(1-\theta)} \right).$$

The parameters of these models can be estimated sequentially by first estimating the conditional probability models $P_{j|i}$, then calculating the index of the aggregate characteristics at the lower level:

$$J_i = \ln\left(\sum_{j'} e^{W_{ij'}} \right)$$

and using this in the [un]conditional probability model P_i. The parameter estimates are consistent although not fully efficient. The standard errors of the estimates must be adjusted to account for the fact that the value J_i is estimated rather than observed.

and using this in the [un]conditional probability model P_i. The parameter estimates are consistent although not fully efficient. The standard errors of the estimates must be adjusted to account for the fact that the value J_i is estimated rather than observed.

The term θ is a measure of the correlation among the error terms. The special case presented here allows correlation among the random components for space heating given air conditioning. If $\theta = 0$, then the error terms are independently distributed and the specification is equivalent to the multinomial logit. Positive values indicate correlation among terms. Values outside the unit interval have not been shown consistent with a model of random utility maximization and imply some counter-intuitive cross-elasticities.

The term J_i is called the "inclusive value" and is used as a measure of the overall utility available to the consumer from all the space heating choices available under a given central cooling choice. Thus, the "inclusive value of space heat given central cooling" and the "inclusive value of space heat given no central cooling" are explanatory variables in the regression of central cooling choice.

A computational error appears to have entered EPRI's calculation of the inclusive values used to produce Table 10. The QUAIL statistical package they used to estimate the nested logit model requires that the inclusive value variables for both alternatives (central cooling and no central cooling) be "expanded" and placed in "alternative 1" and "alternative 2" of a single variable used in the regression of central cooling choice. This expansion causes the inclusive value variables to be doubled in length, and a zero to be placed between each of the original numbers. It appears that this occurred not once, but three times. The result is that the true values appear not in every second position as they should, but rather in every eighth position, as shown in Table II. Thus, the regression was carried out on only about one-fourth the data, and that data was associated with the wrong households.

Table II

Alternative Calculations of Space Heat Inclusive Values

household	alternative	[1]	[2]	[3]	[4]
1	1	MD	0	MD	MD
	2	0	0	0	MD
2	1	0	0	0	-2.3019
	2	0	0	0	-2.5198
3	1	0	0	0	-1.5690
	2	0	0	0	-2.4233
4	1	0	0	0	-1.5697
	2	0	MD	MD	-1.8392
5	1	-2.3358	0	-2.3358	-2.1600
	2	0	0	0	-3.0081
6	1	0	0	0	-2.0550
	2	0	0	0	-2.0490
7	1	0	0	0	-1.1979
	2	0	0	0	-2.0318
8	1	0	0	0	-2.5078
	2	0	-2.5198	-2.5198	-2.7454

[1] = inclusive value given central cooling; overexpanded
[2] = inclusive value given no central cooling; overexpanded
[3] = sum of [1] and [2]; used to get EPRI's Table 10
[4] = corrected; zero-income households excluded
MD = missing data

The different versions of the inclusive value variable in Table II (columns [3] and [4]) produce different estimates of the coefficients of central cooling choice, as shown in Table III.

Table III

Effect of Different Inclusive Values on Central Cooling Choice

variable name	[3] above (incorrect) from Table 10: logit estimate	t-statistic	[4] above; zero-income households excluded: logit estimate	t-statistic
normalized capital cost	-5.849	(-5.369)	-5.152	(-5.076)
normalized operating cost	-181.5	(-5.693)	-126.1	(-3.464)
summer climate	0.1374	(10.05)	0.1396	(10.25)
central cooling choice × income	0.05874	(7.047)	0.07177	(7.605)
inclusive value	0.00164	(0.03410)	0.3013	(2.704)
central cooling choice	-0.09492	(-0.1919)	-1.250	(-2.016)
Likelihood Ratio Test				
log likelihood:	-596.2		-592.6	
log likelihood revised model restricted to original results:			-597.0	
2 × difference ($\approx \chi^2_6$)			8.8 ($p < 0.20$)	

The same likelihood ratio test used previously can test the magnitude of the difference between these two sets of parameter estimates relative to the uncertainty of the "revised" values. Testing a null hypothesis that the true values of the revised parameters are equal to the values reported in EPRI's Table 10, we find that the value of the likelihood ratio statistic is significant at the 20% level. If we test an alternative null hypothesis that only the three parameters exhibiting the greatest change (operating cost, inclusive value, and the central cooling dummy) are equal to EPRI's, we get the same value of the likelihood ratio statistic (8.8), but in this case the statistic is asymptotically distributed as a χ^2_3 random variable. As such, it is significant at the 5% level. On this basis we reject a hypothesis that the true values of those three parameters are equal to the values reported in EPRI's Table 10, and argue instead that the revised values are different than the original ones.

The changes in the coefficients mostly represent "improvements" in the model. The coefficients of operating cost and the inclusive value become less extreme, and the coefficients of inclusive value and central cooling choice become significant, which they are not under the original data set. There is a minor improvement in the quality of the fit, as represented by the percent correctly predicted.

The value of the coefficient of "central cooling choice" (-1.250) indicates a negative utility associated with it, which is somewhat counter-intuitive and must be interpreted in conjunction with the coefficient of "central cooling choice × income". Income is expressed in this data set as multiples of $1000, with the average being about 21.2. Thus, the "average" household will derive positive utility from choosing central air conditioning. In fact, the model suggests that all households with incomes greater than $17.5K derive positive utility from central cooling, and households with incomes less than that derive negative utility from it (i.e., can't afford it, and usually don't select it). This interpretation, given the model, is borne out by repeating the revised regression with all incomes reduced by 17.5 (not shown). The resulting coefficient estimates are unchanged except for the coefficient of central cooling choice, which is essentially zero and not at all significant.

The most important changes are in the coefficients of operating cost and the inclusive value. The change in the operating cost coefficient is important when attempting to estimate market share elasticities from this model, as discussed in the next section. The coefficient of inclusive value is an estimate of the quantity $1 - \theta$. The extremely small (and not significant) value found in the original specification, 0.00164, implies an estimate of θ very close to one, indicating almost perfect correlation among the random components of utility in each branch of the choice hierarchy. The coefficient found under the revised specification, 0.3014, is more credible.

As discussed in EPRI's paper, the relative magnitude of the coefficients of capital and operating cost imply a consumer "discount rate" for the value of money. Assuming static price expectations and very long-lived appliances (so that the effect of depreciation is negligible), then the discount rate is just the ratio of the coefficient of capital cost to that of operating cost. As a result of our revisions, the implied discount rate for the purchase of central cooling increases from 3.2% under the old specification to 4.1% under the revised one. (We are indebted to Dr. Goett of Cambridge Systematics for pointing out this result.) Under both the original and revised models, specifications which would allow the discount rate to vary with income produce estimates of coefficients with counter-intuitive signs (positive when logic tells us they should be negative), and so are rejected.

3. Consequences of Errors (Elasticity Estimates)

We estimate market share elasticities using a sample enumeration approach developed at Lawrence Berkeley Laboratory.[3] The method proceeds by perturbing the value of an independent variable by a fraction δ, averaging the resulting change in market shares over the 1300 households in the database, and dividing by δ and the original (unperturbed) market share. This process produces estimates of the arc elasticities which are approximately quadratic in δ for small perturbations. We take as the point elasticity the intercept of the least-squares regression curve of the arc elasticity estimates as a quadratic function of perturbation size δ, for δ in the range -0.333 to 0.5.

The errors in EPRI's original report have a observable effect on the elasticities calculated from them. Table IV shows the elasticities of eight major fuel/technology categories with respect to own fuel prices, own capital costs, and household income. Elasticities are calculated using both the revised and original versions of Tables 8 and 10 in EPRI's report, with the difference shown in the last column.

Not surprisingly, the elasticities with respect to variables whose coefficients changed the most in revising EPRI's Table 10 also show the most difference here. The coefficients of capital cost, operating cost, and income changed by 13, 44, and 18 percent (as a fraction of the revised value), respectively. Corresponding to these are the elasticities of air conditioning market share with respect to its own capital cost, the price of electricity, and household income. The changes in those elasticities (relative to the revised value) varied from seven to fourteen percent, higher than most of the other changes. Heat pumps, a heating choice available only with central cooling, also shows relatively large changes in elasticities (nineteen percent for elasticity with respect to fuel price). Although the relative change in the coefficient of the inclusive value term was the largest change in EPRI's Table 10 (99%), the effect is felt equally in all the elasticities and so is less detectable.

The changes in EPRI's Table 10 reduced in absolute value the coefficients of two variables (capital cost and operating cost), and increased the coefficient of household income. Therefore it is not surprising that elasticities with respect to those variables generally followed the same pattern: the revised elasticities with respect to capital and operating costs are mostly smaller in absolute value than the original ones, and revised income elasticities are larger.

Despite the noticeable differences in elasticity estimates between the revised and original models, the actual effect on predicted market shares is not large for most economic applications. This is partly due to the major increase in the coefficient of inclusive value. In the original model, the extremely small value of this coefficient meant that very little of the consumer's choice of central cooling was explained by the utility of the space heating alternatives available with the central cooling choice. The larger value of that coefficient in the revised model means that much more of the burden of explanation is being carried by the "lower level" regressions. Since those lower level regressions are essentially unchanged in the revised model (changes in Table 8 being negligible), the effect of the changes in Table 10 is diluted when predicting how market shares change with changes in the independent variables.

The consequences of these elasticity differences are shown in Table V, where the predicted market shares for three different heating/cooling choices are shown under ten, twenty, and fifty percent increases in the price of electricity. Unperturbed market shares are shown in parentheses. The differences in predicted market shares are not particularly significant in the economic sense. Under different perturbations or perturbations in more than one variable (e.g., changes in the capital cost of air conditioning, and/or changes in household income), the difference in predicted market shares could become either greater or lesser, depending on the exact combination of perturbations.

---

[3] Please see the LBL-20090 "Market Share Elasticities for Fuel and Technology Choice in Home Heating and Cooling" for a more detailed discussion of calculating elasticities.

Table IV

Market Share Own-Elasticities for Space Conditioning Equipment

Equipment	Revised Elasticity	Original Elasticity	Difference
	ENERGY COST		
Gas Central Heater	-0.532	-0.534	0.002
Gas Non-Central Heater	-0.226	-0.242	0.016
Oil Central Heater	-1.212	-1.207	-0.005
Oil Non-Central Heater	-0.901	-0.870	-0.031
Electric Central Heater	-1.160	-1.176	0.016
Electric Non-Central Heater	-0.896	-0.854	-0.042
Central Air Conditioner	-0.344	-0.391	0.047
Heat Pump	-0.304	-0.365	0.061
	CAPITAL COST		
Gas Central Heater	-0.377	-0.372	-0.005
Gas Non-Central Heater	-1.502	-1.534	0.032
Oil Central Heater	-1.214	-1.164	-0.050
Oil Non-Central Heater	-1.457	-1.471	0.014
Electric Central Heater	-0.609	-0.611	0.002
Electric Non-Central Heater	-0.915	-0.889	-0.026
Central Air Conditioner	-0.196	-0.216	0.020
Heat Pump	-1.770	-1.709	-0.061
	HOUSEHOLD INCOME		
Gas Central Heater	0.074	0.076	-0.002
Gas Non-Central Heater	-0.575	-0.531	-0.044
Oil Central Heater	-0.298	-0.279	-0.019
Oil Non-Central Heater	-0.590	-0.563	-0.027
Electric Central Heater	-0.146	-0.147	0.001
Electric Non-Central Heater	-0.260	-0.253	-0.007
Central Air Conditioner	0.300	0.276	0.024
Heat Pump	0.385	0.359	0.026

Table V

Predicted Market Shares for Selected Technologies

Change in Price of Electricity:

Technology	+10% Predicted Market Share	+10% Difference	+20% Predicted Market Share	+20% Difference	+50% Predicted Market Share	+50% Difference
Central Gas (0.5121)						
revised:	0.5273		0.5380		0.5585	
original:	0.5277		0.5387		0.5594	
		-0.0004		-0.0007		-0.0009
Central Cooling (0.6414)						
revised:	0.6195		0.5978		0.5335	
original:	0.6160		0.5900		0.5101	
		0.0035		0.0078		0.0234
Heat Pumps (0.1150)						
revised:	0.1112		0.1069		0.0946	
original:	0.1104		0.1052		0.0889	
		0.0008		0.0017		0.0057

4. Extensions and Modifications to the EPRI Model

Heat Pump Cost Calculations

A possible source of error in EPRI's results comes from their lack of "actual" data on capital and operating costs at the household level. Instead, they were obliged to use surrogate data based on the results of MIT's Thermal Load Model [McFadden & Dubin, 1982]. That model estimates the system capacity and energy consumption for each technology using the known structure size and assumed thermal integrity. Those estimates can then be turned into cost estimates by using construction handbooks (for capital costs) and fuel prices (for operating costs).

While this approach is computationally feasible and works well for older technologies, it seems to have been inaccurate for the (then) relatively new heat pump alternative. Average heat pump capital cost (in the EPRI dataset) is more than four times the average central air conditioning capital cost. For some households, the ratio goes as high as nine or ten times air conditioning cost. (Actual heat pump costs are usually in the range of 1.3 to 1.7 times air conditioning costs.)

This discrepancy seems to have come about because the model did not allow for heat pumps to be supplemented with forced-air resistance heating, as is used with almost all heat pump applications in moderate to severe winter climates. Instead, the model simply seems to have specified a larger heat pump. The cost of these large-capacity heat pumps rises quickly, possibly distorting the database.

If that is the case (and, in fact, even if it is not), it is relatively simple to propose a correction to the database. We suggest that a reasonable upper bound on heat pump cost is the cost of a heat pump sufficient to cool the household in the summer season (assumed to be 1.5 times the cost of a conventional air conditioner), **plus** the cost of an electric forced-air furnace sufficient to heat the household in the winter season. Since both of these costs are assumed to be estimated accurately by the procedures used, they are readily available in the EPRI database.

We calculated this upper bound and tested it in the dataset. The average heat pump capital cost was reduced by almost 50%, as almost 98% of the households were in excess of the upper bound.[4] Some further evidence that the origin of the problem was as we have supposed is the relatively high correlation (0.70) between the absolute size of the correction (i.e., the difference between EPRI's specified heat pump capital costs and the proposed upper bound) and the severity of winter climate (measured by heating degree days).

Such a correction to the capital cost also suggests a correction to the operating cost, since forced-air furnaces are generally less efficient than heat pumps. We made this correction using data from EPRI [1985], estimating total degree-hour data for 48 cities across the country. Using this data, we divided the total heating load (degree-hours below 65 °F) into "heat pump load" (degree-hours between 65 °F and 35 °F) and "electric forced-air load" (degree hours below 35 °F). We then assumed that the fraction of the load met by the heat pump could be done at approximately one-half the cost of electric forced-air heat.[5] We estimated the new operating cost for heat pumps in each household (failing the capital cost upper bound) by modifying the electric forced-air operating cost on the basis of the "half-cost heat pump load" calculation. This modification caused the average heat pump operating cost to increase (as would be expected) by some 4%.

We then substituted these modified values of heat pump capital and operating cost into EPRI's dataset and re-estimated the coefficients of EPRI's Table 8. The results of that estimation are shown in the table below.

---

[4] Since 98% of all households violated the "maximum upper bound," it clearly served more to *define* heat pump capital costs than to merely constrain them on the high end.

[5] This assumption amounts to saying that the average heat pump is about twice as efficient for heating as an electric forced-air system (i.e., the coefficient of performance, COP, is 2). This was based on data from Consumer Energy Council of America Research Foundation [1985].

Table VI
Comparison of "Revised" and "Extended" Versions of EPRI's Table 8

Effect of Heat Pump Cost Corrections on Choice of Space Heating (given central cooling)

variable name	Revised: zero-income households excluded — logit estimate	Revised: t-statistic	Extended: heat pump cost corrections — logit estimate	Extended: t-statistic
normalized capital cost	-0.1657	(-6.640)	-0.2807	(-6.668)
normalized operating cost	-2.473	(-4.757)	-3.016	(-5.563)
operating cost × income	-0.04503	(-3.009)	-0.03876	(-2.503)
gas restrictions type 1	-2.428	(-7.598)	-2.466	(-7.680)
gas restrictions type 2	-1.001	(-2.209)	-0.9247	(-1.980)
gas restrictions type 3	-1.793	(-5.999)	-1.812	(-6.047)
heat pump trend	0.04506	(3.855)	0.03010	(2.634)
oil forced-air choice	-1.827	(-8.750)	-1.645	(-7.686)
elec forced-air choice	0.9907	(3.811)	1.209	(4.123)
heat pump choice	-0.3947	(-1.700)	-0.7047	(-3.388)
elec baseboard choice	-1.097	(-4.170)	-1.101	(-3.751)
Likelihood Ratio Test				
log likelihood:	-932.4		-923.6	
log likelihood extended model restricted to revised results:			-964.0	
2 × difference ($\approx \chi^2_{11}$)			40.4 ($p < 0.01$)	

The log-likelihood test on this new set of coefficients (comparing them with the revised coefficients in the left-hand column of the table) shows a χ^2_{11} statistic of 40.4. This value is significant at the 1% level. Since the data on which the extended model is based seem more reasonable with respect to heat pump costs, and the extended model performs significantly better than the model without cost corrections (judged by the improvement of total log-likelihood from -932.4 to -923.6), we prefer the extended model in future calculations.

Cumulative Gas Restrictions

The period covered by the EPRI study (1975-1979) included several major disruptions in the U.S. energy supply. In particular, many gas utilities were obliged to restrict or even completely prohibit new residential gas service at some point in that period. These restrictions had a significant effect on the market for space heating technologies, making it difficult for econometric modeling to separate fuel price and capital cost effects from non-economic effects. EPRI modeled these effects by including binary variables which indicated the presence or absence of three different levels of restrictions for the gas utility serving each household in the year the house was built. The effects of these variables were all found to be statistically significant.

We wished to investigate whether the effects of gas restrictions extended beyond the period during which they were in force. This could occur either because of a psychological mechanism, as consumers and builders remember the prior curtailments and wish to avoid that possibility in the future, or because of the growth of a sales and service infrastructure for alternative technologies during the curtailment.

We tested this idea by extending the definition of the "gas restrictions" variables so that they were no longer binary (absent = 0, present = 1), but could also be fractional if restrictions had been present in years prior to the construction of a particular residence. We desired to preserve the scale of EPRI's original variable (with a range of zero to one) and its significance that the presence of restrictions in the year of construction should be the "worst" effect possible. However, if restrictions had been in place in years prior to construction, we also desired to take into account both the duration of the curtailment and the time elapsed since it was last in effect.

To this end, we redefined EPRI's gas restrictions variables to be sums of negative powers of two, for those households with restrictions in effect in the years prior to construction, but not in effect in the actual year of construction. Thus, one year of restrictions, lifted the year before construction would be scored at 0.5, two consecutive years would score 0.75, two consecutive years lifted two years before construction would score 0.375, etc.

There is nothing logically compelling about this particular arrangement, but it did have the desirable properties of scale and significance discussed above. It also had the property that a (theoretically possible) infinite string of years with gas curtailment, just lifted one year ago, would have the same effect on the installation of gas this year as an outright prohibition this year: both situations would score as 1.0 in the gas restriction variable. Some other progression of values, rather than the geometric one used here, might prove to model these effects more accurately.

Changing the gas restrictions variables in this way significantly improved the quality of the fit of the model to the data, as shown in Tables VII and VIII.

Table VII
Comparison of "Extended" Version of EPRI's Table 8
With and Without "Cumulative" Gas Restrictions

Effect of Cumulative Gas Restrictions on Choice of Space Heating (given central cooling)	Extended: heat pump cost corrections		Further Extended: cumulative gas restrictions	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-0.2807	(-6.668)	-0.2870	(-6.768)
normalized operating cost	-3.016	(-5.563)	-3.324	(-5.790)
operating cost × income	-0.03876	(-2.503)	-0.03713	(-2.340)
gas restrictions type 1	-2.466	(-7.680)	-2.872	(-9.172)
gas restrictions type 2	-0.9247	(-1.980)	-1.111	(-2.604)
gas restrictions type 3	-1.812	(-6.047)	-1.528	(-5.483)
heat pump trend	0.03010	(2.634)	0.02776	(2.407)
oil forced-air choice	-1.645	(-7.686)	-1.828	(-8.518)
elec forced-air choice	1.209	(4.123)	1.182	(4.011)
heat pump choice	-0.7047	(-3.388)	-0.8220	(-3.920)
elec baseboard choice	-1.101	(-3.751)	-1.158	(-3.927)
Likelihood Ratio Test				
log likelihood:	-923.6		-898.3	
log likelihood "cumulative" model restricted to "non-cumulative" results:			-902.7	
2 × difference ($\approx \chi^2_{11}$)			8.8	

Table VIII
Comparison of EPRI's Table 5
With and Without "Cumulative" Gas Restrictions

Effect of Cumulative Gas Restrictions on Choice of Space Heating (given no central cooling)	EPRI's Table 5: original coefficients		Extended: cumulative gas restrictions	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-0.4843	(-4.683)	-0.5069	(-4.738)
normalized operating cost	-2.259	(-5.773)	-2.234	(-5.939)
gas restrictions type 1	-2.715	(-5.871)	-3.255	(-7.150)
gas restrictions type 2	-1.400	(-3.477)	-1.625	(-4.286)
gas restrictions type 3	-1.033	(-3.359)	-0.5741	(-1.933)
gas hydronic choice	-1.826	(-3.779)	-1.727	(-3.525)
gas non-central choice	-3.562	(-8.550)	-3.530	(-8.463)
oil forced-air choice	-1.001	(-4.456)	-1.250	(-5.286)
oil hydronic choice	0.09985	(0.2523)	-0.09038	(-0.2202)
oil non-central air choice	-3.976	(-6.672)	-4.165	(-6.950)
elec forced-air choice	-1.187	(-3.746)	-1.403	(-4.482)
elec baseboard choice	-1.183	(-3.946)	-1.397	(-4.673)
Likelihood Ratio Test				
log likelihood:	-560.0		-533.3	
log likelihood "cumulative" model restricted to original results:			-538.0	
2 $\times$ difference ($\approx \chi^2_{12}$)			9.4	

This table shows a dramatic improvement (from -560.0 to -533.3) in the log likelihood when shifting to cumulative gas restrictions. In fact, some of this improvement, approximately 10 "points" worth, is due to the fact that the model with cumulative gas restrictions is run on fewer households than the original model (nine households fewer, to be exact). This slight reduction in the size of the dataset was necessary because of mismatches between different variables in different parts of the EPRI data, and produces a necessarily better (i.e., lower) overall log likelihood. In fact, the original EPRI model, without cumulative gas restrictions, has an overall log likelihood of -550.2 on the reduced dataset.

As can be seen by examination of Tables VII and VIII, the proposed change produces a significant improvement in the overall log-likelihood of each submodel (improvement from -923.6 to -898.3, and from -550.2 to -533.3, respectively), even though the coefficients with cumulative gas restrictions are not significantly different than those without (when judged on the "cumulative" data set).

Table IX shows the results of calculating inclusive values from these new "cumulative" coefficients (in Tables VII and VIII) and recalculating the upper-level regression (i.e., choice of central cooling or no central cooling) shown in EPRI's Table 10. Since the proposed changes in variable definitions (heat pump cost corrections and cumulative gas restrictions) occur only at the "lower level", they affect this regression only through the inclusive value term. Therefore, it is not surprising that there is only a very slight improvement in the overall log likelihood (from -592.6 to -592.4) of this regression. Nonetheless, these changes represent a slight improvement at this level of the model and significant improvements at the lower level, and we feel that further analysis and use of the EPRI data should take them into account.

Table IX
Comparison of EPRI's Table 10
With and Without Heat Pump Cost Corrections
and "Cumulative" Gas Restrictions

Effect of Proposed Corrections and Extensions on Choice of Central Cooling	EPRI's Table 10: corrected coefficients from Table III		Extended: heat pump corrections and cumulative gas restrictions	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-5.152	(-5.076)	-4.801	(-4.895)
normalized operating cost	-126.1	(-3.464)	-128.0	(-3.505)
summer climate	0.1396	(10.25)	0.1433	(10.48)
central cooling choice × income	0.07177	(7.605)	0.07195	(7.528)
inclusive value	0.3013	(2.704)	0.4028	(2.791)
central cooling choice	-1.250	(-2.016)	-1.155	(-1.988)
Likelihood Ratio Test				
log likelihood:	-592.6		-592.4	
log likelihood revised model restricted to original results:			-596.9	
2 × difference ($\approx \chi^2_6$)			9.0 ($p < 0.20$)	

Price Expectations

In this section, we report an "unsuccessful" effort to extend the EPRI model to account for possible consumer expectations regarding fuel prices. The logic for this is that consumers can reasonably be presumed to have such expectations, and may be taking them into account when selecting their space heating alternative. Modeling these expectations, however, is not as simple a matter as we might hope.

Some authors [Hartman and Hollyer, 1977, and Dohrmann, 1980] have approached this problem by including actual fuel prices from years in the "future" for the consumer's decision (but, obviously, in the past from the point of view of the researcher). In effect, such a model proposes that consumers can predict future fuel price (or at least, future trends in fuel prices) accurately, and act accordingly. Although recent history makes this seem unlikely, such a model works quite well if prices and price inflation are reasonably stable over the modeled period.

We chose a slightly different approach of assuming that decision-makers were aware of the actual change in prices for each fuel in the year prior to construction of their residence, and that they extrapolated that growth rate for an unspecified number of years when considering the operating cost of each alternative.[6]

We then recalculated the regression models for several different versions of this price variable: one, five, and ten year linear and compounded growth rates, and non-negative versions of these (in which any decline in actual prices is presumed to be disbelieved, and the growth rate is modeled as zero) . All of them produced poorer fits to the dependent variable (judged by the overall log likelihood of the regression) than the corrected and extended version of EPRI's models shown in Tables VII, VIII, and IX.[7] The least decline in predictive power was for the one-year, nonegative version of the price expectations. The results for these models are shown in Tables X, XI, and XII. We would like to emphasize that we are **not** recommending any further use of the coefficients shown in those tables; they are provided only to demonstrate the decline in overall fit of the model. This decline does not conclusively rule out the value of information about future price expectations in consumer decision modeling. But it does suggest that the effects are more complex than generally believed, and that the assumptions underlying any model of them should be examined carefully.

[6] We chose this approach at least partially because of the ready availability of the necessary data in EPRI's dataset: fuel prices in the year prior to construction were available for each household. Later conversations with Dr. Andrew Goett, the principal author of the EPRI report, indicated that these data were available precisely because they had considered, and rejected, a similar approach to modeling price expectations.

[7] A similar effect occurred when the price expectations models were tried with EPRI's original models (i.e., without the heat pump cost corrections or cumulative gas restrictions).

Table X
Comparison of Extended Version of EPRI's Table 8
With and Without "Price Expectations"

Effect of Price Expectations on Choice of Space Heating (given central cooling)				
	From Table VII: heat pump corrections & cum gas restrictions		Further Extended: fuel price expectations	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-0.2870	(-6.768)	-0.2809	(-6.599)
normalized operating cost	-3.324	(-5.790)	-1.511	(-3.599)
operating cost × income	-0.03713	(-2.340)	-0.03383	(-2.421)
gas restrictions type 1	-2.872	(-9.172)	-2.720	(-8.592)
gas restrictions type 2	-1.111	(-2.604)	-0.9240	(-2.209)
gas restrictions type 3	-1.528	(-5.483)	-1.373	(-4.986)
heat pump trend	0.02776	(2.407)	0.03444	(3.010)
oil forced-air choice	-1.828	(-8.518)	-2.173	(-10.57)
elec forced-air choice	1.182	(4.011)	0.1854	(0.8071)
heat pump choice	-0.8220	(-3.920)	-1.120	(-5.562)
elec baseboard choice	-1.158	(-3.927)	-1.951	(-7.613)
Likelihood Ratio Test				
log likelihood:	-898.3		-917.4	
log likelihood "Table VII" model restricted to "price expectations" results:	-909.7			
2 × difference ($\approx \chi^2_{11}$)	22.8 ($p < 0.02$)			

Table XI
Comparison of Extended Version of EPRI's Table 5
With and Without "Price Expectations"

Effect of Price Expectations on Choice of Space Heating (given no central cooling)				
	From Table VIII: cumulative gas restrictions		Further Extended: fuel price expectations	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-0.5069	(-4.738)	-0.6166	(-5.757)
normalized operating cost	-2.234	(-5.939)	-1.325	(-5.115)
gas restrictions type 1	-3.255	(-7.150)	-3.188	(-7.008)
gas restrictions type 2	-1.625	(-4.286)	-1.533	(-4.055)
gas restrictions type 3	-0.5741	(-1.933)	-0.6125	(-2.135)
gas hydronic choice	-1.727	(-3.525)	-1.363	(-2.831)
gas non-central choice	-3.530	(-8.463)	-3.456	(-8.294)
oil forced-air choice	-1.250	(-5.286)	-1.270	(-5.440)
oil hydronic choice	-0.09038	(-0.2202)	0.1928	(0.4726)
oil non-central air choice	-4.165	(-6.950)	-4.187	(-6.999)
elec forced-air choice	-1.403	(-4.482)	-1.885	(-6.994)
elec baseboard choice	-1.397	(-4.673)	-1.836	(-6.952)
Likelihood Ratio Test				
log likelihood:	-533.3		-542.1	
log likelihood "Table VIII" model restricted to "price expectations" results:	-536.9			
2 × difference ($\approx \chi^2_{12}$)	12.6 ($p < 0.05$)			

Table XII
Comparison of Corrected Version of EPRI's Table 10
With and Without Price Expectations

Effect of Price Expectations on Choice of Central Cooling	From Table IX: corrected coefficients		Extended: fuel price expectations	
variable name	logit estimate	t-statistic	logit estimate	t-statistic
normalized capital cost	-4.801	(-4.895)	-4.148	(-4.323)
normalized operating cost	-128.0	(-3.505)	-51.76	(-2.033)
summer climate	0.1433	(10.48)	0.1431	(10.54)
central cooling choice × income	0.07195	(7.528)	0.07004	(7.531)
inclusive value	0.4028	(2.791)	0.3772	(3.174)
central cooling choice	-1.155	(-1.988)	-2.160	(-3.863)
Likelihood Ratio Test				
log likelihood:	-592.4		-597.8	
log likelihood "Table IX" model restricted to "price expectations" results:	-601.6			
2 × difference ($\approx \chi^2_8$)	18.4 ($p < 0.01$)			

Discount Rates Varying with Household Characteristics

In models using a linear combination of exogenous variables to estimate consumer utility, the ratio of the capital cost coefficient to operating cost coefficient gives an estimate of an implicit discount rate for the choices reported in that data set.[8] Standard economic theory suggests that this discount rate should fall with increasing income, and this result may be observed by including the product of operating cost and income as an exogenous variable of the model. EPRI did this in both of the two heating choice sub-models, finding the expected result in one case and a counterintuitive result (discount rates rising with increasing income) in the other. In that case, the formulation was rejected in favor of using only operating cost (and not operating cost times income) as an exogenous variable. This simpler formulation in "pure" capital and operating cost effects implies that the decisions represented in that portion of the data are best modeled by a single fixed discount rate.

We chose to ask whether the very specific discount rates in this problem (associated with the choice of heating and cooling technologies) could profitably be modeled as varying with other household or geographic conditions, rather than just income. In particular, we looked at maximum heat load and summer climate as variables which capture a sense of the desirability and intensity of use of heating and cooling, respectively.

One extension, then, was to include the product of operating cost and maximum heat load as an independent variable. Heat load is positively correlated with income (higher income households tend to live in larger houses, which leads to higher maximum heat loads), and both capital and operating costs (larger houses again). For the population with central cooling, we found that formulations which included operating cost and heat load consistently out-performed (in the statistical sense) sub-models using operating cost and income. Essentially, the data indicate that households with large heating needs (either because of their size, or their geographic location) tend to pay more attention to considerations of future operating cost than do households with small heating loads, irregardless of income.

We also found a relatively simple model, including only "operating cost times heat load," to be only marginally inferior (in overall log-likelihood) to a more complex model which also included a pure "operating cost" term.[9] A comparison of three different formulations (the original, with both operating cost and the product of operating cost times heat load, and with the product term alone) can be seen in Table XIII.

For households without central cooling, we did not find the result as strongly as we expected. There, a formulation with a pure "operating cost only" effect were statistically superior to one with a product "operating cost times heat load" effect only. That formulation, in which household discount rates will vary with the heat load, is not rejected by the data (i.e., the coefficient has the correct sign), but it cannot explain the observed choices as well as a constant discount rate across all households.

When we included both exogenous variables in the same model, the coefficient of the product term comes out with a counterintuitive (i.e., positive) sign. This is not completely surprising, given that EPRI obtained the same counterintuitive result for a formulation with both "operating cost" and "operating cost times income" in this portion of the data, and the variables of household income and heat load have a positive overall colinearity ($r = 0.44$).[10] Comparisons of these three alternative formulations can be seen in Table XIV.

[8] See EPRI [1984] for a discussion of this point.

[9] The simpler model (with one fewer exogenous variables) is superior by the Akaike Information Criterion.

[10] Both maximum heat load and income are measured in units of 1000 (Btu/hr and 1975 dollars, respectively).

Table XIII
Comparison of Extended Version of EPRI's Table 8
With and Without Maximum Heat Load

Effect of Maximum Heat Load on Choice of Space Heating
(given central cooling)

	From Table VII: heat pump corrections & cum gas restrictions	Further Extended: discount rate varying w/ maximum heat load	
	op cost × income only	op cost and op cost × heat load	op cost × heat load only
variable name	logit estimate	logit estimate	logit estimate
normalized capital cost	-0.2870	-0.1971	-0.1855
normalized operating cost	-3.324	-0.5919*	
operating cost × income	-0.03713		
operating cost × heat load		-0.07442	-0.08264
gas restrictions type 1	-2.872	-3.016	-3.022
gas restrictions type 2	-1.111	-1.532	-1.560
gas restrictions type 3	-1.528	-1.724	-1.732
heat pump trend	0.02776	0.03187	0.03293
oil forced-air choice	-1.828	-2.022	-2.079
elec forced-air choice	1.182	1.074	0.9305
heat pump choice	-0.8220	-0.9822	-1.039
elec baseboard choice	-1.158	-1.157	-1.259
overall log likelihood:	-898.3	-888.8	-889.1

* This coefficient has a t-statistic such that the estimate is not significantly different from zero.

Table XIV
Comparison of Extended Version of EPRI's Table 5
With and Without Maximum Heat Load

Effect of Maximum Heat Load on Choice of Space Heating (given no central cooling)			
	From Table VIII: cumulative gas restrictions	Further Extended: discount rate varying w/ maximum heat load	
	op cost only	op cost and op cost × heat load	op cost × heat load only
variable name	logit estimate	logit estimate	logit estimate
normalized capital cost	-0.5069	-0.5288	-0.5033
normalized operating cost	-2.234	-2.703	
operating cost × heat load		0.01163*	-0.04235
gas restrictions type 1	-3.255	-3.243	-3.302
gas restrictions type 2	-1.625	-1.586	-1.732
gas restrictions type 3	-0.5741	-0.5672	-0.6199
gas hydronic choice	-1.727	-1.656	-1.730
gas non-central choice	-3.530	-3.521	-3.514
oil forced-air choice	-1.250	-1.221	-1.369
oil hydronic choice	-0.09038	-0.004721	-0.2052
oil non-central air choice	-4.165	-4.141	-4.271
elec forced-air choice	-1.403	-1.441	-1.646
elec baseboard choice	-1.397	-1.436	-1.598
overall log likelihood:	-533.3	-533.1	-539.5

* This coefficient is counterintuitive (i.e., has positive sign when economic intuition predicts a negative sign), and has a t-statistic such that the estimate is not significantly different from zero.

The partial success of adding maximum heat load to discount rate considerations in the heating choice sub-models prompted us to consider a formulation of the central cooling choice model with summer climate playing the same role.[11] Thus, we tried out a formulation with a product of operating cost and summer climate as an exogenous variable (as well as a pure climate effect) and found it statistically superior (measured by the overall log-likelihood) to the original model. A comparison of the two models is shown in Table XV.

[11] EPRI defined an unusual variable to model summer climate, which appears to be highly successful. Noting that the traditional measure of summer climate (cooling degree days) fails to adequately capture personal discomfort (or at least, personal discomfort is highly non-linear in cooling degree days), EPRI defined a new variable as the dry-bulb summer design temperature minus 82 °F plus the wet-bulb summer design temperature minus 68 °F, with a minimum of zero. This variable appears to capture the relevant aspects of perceived need for air conditioning more completely than does cooling degree days.

Table XV
Comparison of Corrected Version of EPRI's Table 10
With and Without Summer Climate

	Effect of Summer Climate on Choice of Central Cooling		
	From Table IX: corrected coefficients	Extended: inclusive value from Tables XIII and XIV	Further Extended: discount rate varying w/ summer climate
variable name	logit estimate	logit estimate	logit estimate
normalized capital cost	-4.801	-4.779	-3.585
normalized operating cost	-128.0	-153.8	
operating cost × summer climate			-11.04
summer climate	0.1433	0.1445	0.2518
central cooling choice × income	0.07195	0.06435	0.06456
inclusive value	0.4028	0.1157**	0.1362**
central cooling choice	-1.155	-0.7872	-2.461
overall log likelihood:	-592.4	-597.5	-589.9

We feel that these results of discount rates varying with heating and cooling characteristics have a strong intuitive appeal, and we expect to be able to replicate these results in future econometric studies. Essentially, the data support the notion that households for which heating and cooling costs are likely to be large (i.e., those households with a large heat load or a summer climate which makes air conditioning highly desirable) are likely to pay more attention to the tradeoff between future operating costs and immediate capital costs in their choice of equipment.

We note that one consequence of our preferred formulations in Tables XIII and XV is that a household with zero heating load or "zero" summer climate is predicted to have an infinite discount rate for the choice of heating or cooling equipment, respectively. While we are uncomfortable when any economic factor is modeled as going to infinity, we feel that the model does approximate the truth: a household with no heating need should reasonably be unwilling to pay any additional amount of capital cost to reduce future operating costs, since anticipated operating costs are already zero.[12]

We also note that higher values of maximum heating load and the summer climate variable are both associated with increased operating costs, and our inclusion of those variables as products with operating cost implies that consumer utility seems to be linear in an "operating cost times operating cost" effect. However, all formulations which included a squared operating cost variable directly (and similar products or squares of the capital cost term) turned out to be statistically

** In both formulations shown, the inclusive value is based on our choice for the overall best model of the two heating-choice submodels in Tables XIII and XIV (i.e., using operating cost times heat load in the choice of heating with central cooling, and operating cost alone in the choice of heating without central cooling.

[12] The fact that we do not observe a similar effect in households without air conditioning is at least slightly bothersome. One rationalization is that it may turn out that the experience of anticipating future cooling costs and making the appropriate tradeoffs of capital and operating costs is what sensitizes a homeowner to all space conditioning costs. By this notion, households without the need for air conditioning simply receive less exposure to issues of future vs. present costs, and so fail to consider their heating load to a significant degree. Remember, for the heating choice sub-model without central cooling, a formulation with discount rates that varied by heating load was not rejected by the data. It was merely supported less strongly than a model with constant discount rates.

inferior to those presented here.

Lastly, we note that we have not made all the same statistical comparisons in Tables XIII, XIV, and XV as we did in all earlier tables. In Tables I through XII, we were redefining some of the existing variables and sought to test whether the old coefficients (of the redefined variables) were significantly different from the new ones. In this last section (Tables XIII through XV), we are exchanging new exogenous variables for old, making the particular χ^2 test used earlier less appropriate. A direct comparison of overall log-likelihoods is still a reasonable test of the relative superiority of two different models, however.

5. Conclusions

This paper has examined errors in and possible extensions to the models of consumer choice of space heating appliances reported in EPRI EA-3409. Two minor calculation errors in EPRI's report were discussed in Section 2 of this paper. In Section 3, we examined the significance of these errors and found them to be not particularly serious.

We also consider four possible extensions to EPRI's work, reported in Section 4. Three of these four extensions (correcting heat pump costs, modeling gas restrictions as cumulative in effect, and discount rates varying with household characteristics) significantly improve the fit of the model to the data, and we suggest that coefficients from the extended models in Tables XIII, IX, and XV be used in place of the original coefficients. The fourth proposed extension (including a "price expectations" effect) did not improve the model, and despite its theoretical attractiveness, we do not feel it can be adequately modeled in this dataset.

References

Consumer Energy Research Council of America Research Foundation [1985]: "Heat Pumps Versus Alternative Space Conditioning Systems: A Model to Evaluate Life Cycle Costs with Site-Specific Applications," September, 1985, page 28.

Dohrmann, D. [1980]: "Estimation of Fuel Choice Elasticities for the Residential Sector," report prepared for Lawrence Berkeley Laboratories.

Electric Power Research Institute [1984]: "Household Appliance Choice: Revision of REEPS Behavioral Models," EPRI EA-3409, Research Project 1918-1, Final Report.

Electric Power Research Institute [1985]: "Heat Pump Manual" EPRI EM-4110-SR, Special Report, August, page 70.

Hartman, R. and Hollyer, M. [1977]: "An Examination of the Use of Probability Modeling for the Analysis of Interfuel Substitution in Residential Energy Demand," MIT Energy Laboratory Working Paper MIT-EL-77-018WP, July.

McFadden, D. [1978]: "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models,* edited by Karlqvist, A. *et al.,* Amsterdam: North-Holland.

McFadden, D. and Dubin, J. [1982]: "A Thermal Model for Single Family Owner-Occupied Detached Dwellings in the National Interim Energy Consumption Survey," MIT Energy Laboratory Discussion Paper No. 25, MIT-EL-82-040WP, June.

Wood, D., Ruderman, H., and McMahon, J. [1989]: "Market Share Elasticities for Fuel and Technology Choice in Home Heating and Cooling," LBL Technical Report LBL-20090.

Appendix

The text of this report contains a number of comparisons of alternative sets of coefficients for particular regressions. In each case, the different sets of coefficients arise due to different proposals for the definitions or treatment of certain variables. And in each case, we can designate one set of coefficients as arising from a "revised" specification of the model.[13]

Both the "original" and "revised" coefficients presented throughout this paper are estimates of "true" coefficients that would be obtained if 1) the nested logit model is correct (i.e., all the assumptions of the model are met, and the particular specification of the variables used is correct); and 2) if we could obtain a complete sampling of the population. As estimates, they each suffer from some uncertainty, having asymptotic normal distributions. We desire a test that would tell us just how significant the difference between the two estimates is, given the inherent uncertainty of both estimates.

Let $\hat{\beta}$ = the estimated parameter vector from the "original" regression,
$\hat{\alpha}$ = the estimated parameter vector from the "revised" regression,
Σ = the variance of their difference, i.e., $\mathrm{Var}(\hat{\beta} - \hat{\alpha})$,
$= \mathrm{Var}(\hat{\beta}) + \mathrm{Var}(\hat{\alpha}) - 2\mathrm{Cov}(\hat{\beta},\hat{\alpha})$.

Then

$$d = (\hat{\beta} - \hat{\alpha})^T \Sigma^{-1} (\hat{\beta} - \hat{\alpha})$$

will be distributed approximately as a χ^2 random variable with degrees of freedom equal to the number of parameters estimated.

Unfortunately, we are lacking an estimate of the term $\mathrm{Cov}(\hat{\beta},\hat{\alpha})$, although it seems reasonable to presume that, as a covariance matrix, it is positive definite and probably something like the $\mathrm{Var}(\hat{\beta})$ or $\mathrm{Var}(\hat{\alpha})$ matrices. This supposition is based on the observation that, in most cases, the data matrices of independent variables from which the coefficients $\hat{\beta}$ and $\hat{\alpha}$ were estimated differed in only one, or at most, a few variables (e.g., the variable "inclusive value", in EPRI's Table 10, discussed in section 2 of this paper). This guesswork will allow us to apply the logistic ratio test and hope that it will understate the real significance of the difference between the two estimated parameter vectors.

The logistic ratio test examines a null hypothesis that the "true" parameters are equal to some set of constants. Our estimated parameters have some uncertainty, and the logistic ratio test asks how probable it is that we should have obtained estimates that differ from the hypothesized values, if the hypothesis is true. If it is unlikely that we would have obtained the values we did get, then it is unlikely that the hypothesis is true. The test is carried out by comparing the value of the log likelihood function at the value of the parameter vector we obtained (which maximizes the likelihood) and at the hypothesized values. Twice the difference in log likelihoods is distributed approximately as a χ^2 random variable with degrees of freedom equal to the number of parameters being compared. Thus the value of the statistic which is found can be compared with tabulated values of the χ^2 distribution to determine its significance.

We can apply the test to compare two alternative sets of coefficients by using the preferred model specification (where preference is based on theoretical attractiveness of the model, and/or the value of the overall log likelihood), and hypothesize that the "true" values are those obtained from the "non-preferred" specification.

In general, such a test is not perfectly adequate for our situation because it ignores the uncertainty in the coefficients of the alternative specification, which we set equal to the (constant) hypothesized values of the test. Only the variability in the preferred estimates is admitted to.

[13] Examples include the alternative treatment of zero-income households in EPRI's Table 8 (discussed in section 2 of this paper), or the cumulative versus non-cumulative definition of gas restrictions (discussed in section 4).

Essentially we are calculating:

$$d = (\hat{\beta} - \hat{\alpha})^T \left[\mathrm{Var}(\hat{\alpha})\right]^{-1} (\hat{\beta} - \hat{\alpha})$$

instead of the correct statistic, d, above.

The question then arises, under what conditions will the logistic ratio test under- or overstate the significance of the difference between the two estimated parameter vectors? By analogy with the one-dimensional case, we can show that the likelihood ratio statistic will understate the significance when

$$x^T \left[2\mathrm{Cov}(\hat{\beta},\hat{\alpha}) - \mathrm{Var}(\hat{\beta})\right] x > 0 \qquad \text{for all } x,$$

that is, $[2\mathrm{Cov}(\hat{\beta},\hat{\alpha}) - \mathrm{Var}(\hat{\beta})]$ is positive definite. If we are correct in our belief that $\mathrm{Cov}(\hat{\beta},\hat{\alpha})$ is something like $\mathrm{Var}(\hat{\beta})$, then the sufficient conditions for understating the significance will probably be met.

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
1 CYCLOTRON ROAD
BERKELEY, CALIFORNIA 94720