Refrigerator/Freezer UEC Estimation, 1996
ARCA/SCE Turn-in Program

(In Support of XENERGY Inc.'s Evaluation of the 1996 Appliance Recycling Program)

SCE Study #537

John Peterson
Athens Research

May 1, 1998

I. INTRODUCTION: BACKGROUND AND SUMMARY OF FINDINGS.

This study reports on regression-based estimates of the annual kWh consumption of refrigerators and freezers that were collected in Southern California Edison's 1996 appliance turn-in program, a program administered by the Appliance Recycling Centers of America (ARCA). The study combines tracking system data for the 1996 program, DOE-protocol metering from a recent study conducted by ARCA (hereafter "ARCA metering study"), and DOE-protocol metering from an Edison-sponsored study of appliances collected by ARCA in the 1998 version of the recycling program (metering conducted by BR Labs; hereafter "BR metering study"). The ARCA metering study consists of 1173 appliances sampled from various ARCA sites and programs in the U.S. The BR metering study, in which 140 appliances of an intended 150 were metered and recorded as of late April, 1998, is used to assess and correct for biases that may have emerged from either sample selection or instrumentation in the ARCA metering study. The BR study, as indicated in each of the firm's reports on individual appliances, tested the refrigerators "in accordance with the procedure outlined in Book 10 of the Code of Federal Regulations (CFR), Section 430.23(b), 1997," which is commonly referred to as the "DOE test." The two metering study samples are combined, and a regression including terms assessing/correcting for ARCA metering study biases, is calibrated against metered consumption. This regression, which relates appliance configuration, size, amperage, and age to metered consumption, provides the parameters to which 1996 tracking system population data are applied, to obtain expected metered consumption values for that population (overall, by appliance type, and by configuration/size). The results of this effort are the following estimated average annual kWh (the full year kWh consumption expected for the 1996 program appliance turn-ins; see Table 1).

Table 1: UEC Estimates Based on Two-sample Regression Analysis and Evaluation at 1996 Population Means

	Annual kWh Mean	Standard Error
Refrigerators	2147.6	80.0
Freezers	2058.3	125.9
Overall	2130.1	69.7

These estimates are used by XENERGY, Inc., in its more comprehensive estimation of gross and net 1996 program impacts. The estimates' strong face, inferential, and construct validity are due to the combination of population information and sample-based parameters relating these variables to consumption, as well as the supplementing of the large ARCA metering sample with "bias-checking" data points from the BR sample.

The development and diagnostic review of the regression model used to estimate impacts is included in this report, with a view to satisfying relevant portions of the California DSM Measurement Protocols' Table 7 quality assurance requirements. Additionally, we include preliminary analyses of the extent to which the BR metering study and the ARCA metering study appear to be drawn from the same population, with regard to relevant and available variables. We also consider the extent to which the combined sample adequately represents the 1996 SCE recycled appliance population. Note that with regression terms appropriately specifying the two-sample character of the model as well as a sufficiently wide range of consumption-relevant appliance characteristics, the application of population data to the resultant model removes most concerns about representation.

The following outlines the remaining sections of this report.

II. DATA SOURCES.
- A. Tracking Data
- B. ARCA Metering Study Sample
- C. BR or Current Metering Sample
- D. Summary: Combined Use of 1996 Tracking, ARCA Sample, and BR Sample.

II. DATA SOURCES.

A. Tracking Data.

From the overall 1996 program file, refrigerators or freezers with both 1996 pickup dates and status codes indicating working appliances were extracted for analysis. These 25,407 appliance records were examined for the presence of amperage, size, appliance type, legitimate age range (nonmissing and less than 50 years), and defrost type identification. Winnowing out 73 appliance records bereft of any of this information, a total of 25,334 appliances (20,333 refrigerators and 5,001 freezers) were stored in files POP96 and further recoded file POP96A. Data available in POP96A include appliance type (refrigerator or freezer), defrost type, configuration type, volume, amperage, and age. Other variables not participating in this analysis included appliance brand, location at the dwelling, color, a pickup identifier number, customer and location identifiers, pickup scheduling information, contact information, and various survey items administered by ARCA. Additional terms, mirroring variables created for the sample regressions, were also created in POP96A, in order to allow the application of population means to calibrated regressions to proceed as a relatively smooth post-processing exercise after calibration. Appendix A contains distributional information for the population file.

B. ARCA Metering Study Sample.

The ARCA metering study sample is the same sample of 1173 metered appliances that has been used in previous Xenergy evaluation of the 1994-1995 version of the SCE turn-in program. The refrigerator UEC used by Xenergy came from a Barakat and Chamberlin regression study of the refrigerators in the sample, while the freezer UEC used was based on an Athens Research regression analysis of the sample's freezers. The sample is an amalgam of DOE-metered appliances from several different ARCA recycling centers in the U.S. Variables available include brand, age, appliance type, configuration, defrost type, rated amperage, and metered annual kWh. Distributional information on the ARCA metering sample is given in Appendix B. The ARCA metering study sample's auspices and validity have been questioned by ORA reviewers, so that it is being combined with an independent metering sample that will serve to identify and adjust for alleged biases.

C. BR or Current Metering Sample.

In order to precisely estimate UEC's for the 1996 program, and in the process determine the direction and magnitude of the combined effects of sample and instrumentation bias alleged against the ARCA sample, Edison undertook a collaboration with BR Labs of Huntington Beach to select and meter a sample of 150 appliances during December 1997 and the early months of 1998, under DOE protocols. The collaboration maintained control of the selection process by virtue of Edison/BR's random selection of ARCA collection days and random selection from the shipping list of appliances to fill cells of a sampling plan developed by Athens Research and SCE staff.

The design of the BR sample began with the notion that SCE should be able to see a range of allocation rules for the 150 sample points, and the likely precision results, prior to choosing a particular design. The general intent was that the sample (1) be capable of "standing on its own" – i.e., produce near-optimal precision levels if a simple weighted mean were produced from the 150, (2) allow reasonable power to detect bias in a multivariate context, (3) provide enough variability in configuration and size to allow unbiased estimates of the impacts of all coefficients in such a regression – particularly crucial in that SCE will use regression results to fill a variety of "cells" with UEC estimates for planning purposes. A January 5, 1998 memorandum (Appendix C of this document) describes six distinct allocation plans based on various ways of prioritizing and implementing the three considerations just mentioned. The design finally selected (described as "allocation n5" in the memorandum) allowed for 115 refrigerators and 35 freezers, with a Neyman allocation of sample points to configurations and size groups using annual kWh variances obtained from the "old" ARCA study sample). The allocation is described in the January 5 memorandum in some detail, and summarized below in Table 2. This table shows the original stratification for the plan, which involved unique combinations of appliance type (freezer/refrigerator), defrost type/configuration (frost free, partial defrost, manual defrost; chest freezer, upright freezer, bottom freezer refrigerator, side-side style refrigerator, top freezer refrigerator, single door refrigerator), and size group. Size groupings were tailored somewhat, as described in the original sampling memorandum, to reflect the population distributions within specific configurations. It should be noted that this is the target allocation selected by SCE Measurement and Evaluation, but that only 140 of the target were collected and metered as of April 24, 1998, requiring that the UEC analysis be carried out on this slightly reduced sample, and requiring some collapsing in the coding of regression terms that might otherwise have supported more detail. Four of these 140 were metered by BR but at the time of this analysis we were unable to retrieve their tracking system information on size, configuration, estimated age, etc., and they are therefore not included in this analysis. It is anticipated that a more comprehensive UEC analysis will be completed later this year, using a complement of 150 appliances (i.e,, the four appliances just mentioned, plus ten appliances being metered as this document is written).

Data available in the BR sample include identifier numbers, appliance type, configuration, defrost type, volume, age, and amperage. Appendix D contains unweighted descriptive data on the BR sample.

Table 2: Summary of Planned Sample Allocation, BR Sample

FREEZR/ REFRIG	CONFIGR1	SIZEGRP2	ALLOCATION
F	FF_UF	a: 10-17 f	3
F	FF_UF	b: 18+ ft	5
F	MD_CF	a: 10-13 f	2
F	MD_CF	b 14-17 f	3
F	MD_CF	c: 18-20 f	3
F	MD_CF	d: 21+ ft	2
F	MD_UF	a: 10-17 f	9
F	MD_UF	b: 18+ ft	10
R	FF_BF	a: 10-17 f	5
R	FF_BF	b: 18-20 f	5
R	FF_BF	c: 21+ ft	2
R	FF_SS	a: 14-20 f	4
R	FF_SS	b: 21+ ft	16
R	FF_TF	a: 10-13 f	2
R	FF_TF	b: 14-17 f	11
R	FF_TF	c: 18-20 f	36
R	FF_TF	d: 21+ ft	3
R	MD_SD	a: 10-13 f	3
R	MD_SD	b: 14+ ft	6
R	PD_BF	a: 10-17 f	3
R	PD_BF	b: 18+ ft	3
R	PD_SD	a: 10-17 f	2
R	PD_SS	a: 14-17 f	2
R	PD_SS	b: 18+ ft	2
R	PD_TF	a: 10-13 f	2
R	PD_TF	b: 14-17 f	4
R	PD_TF	c: 18+ ft	2
			=====
			150

D. Summary: Combined Use of 1996 Tracking, ARCA Sample, and BR Sample.

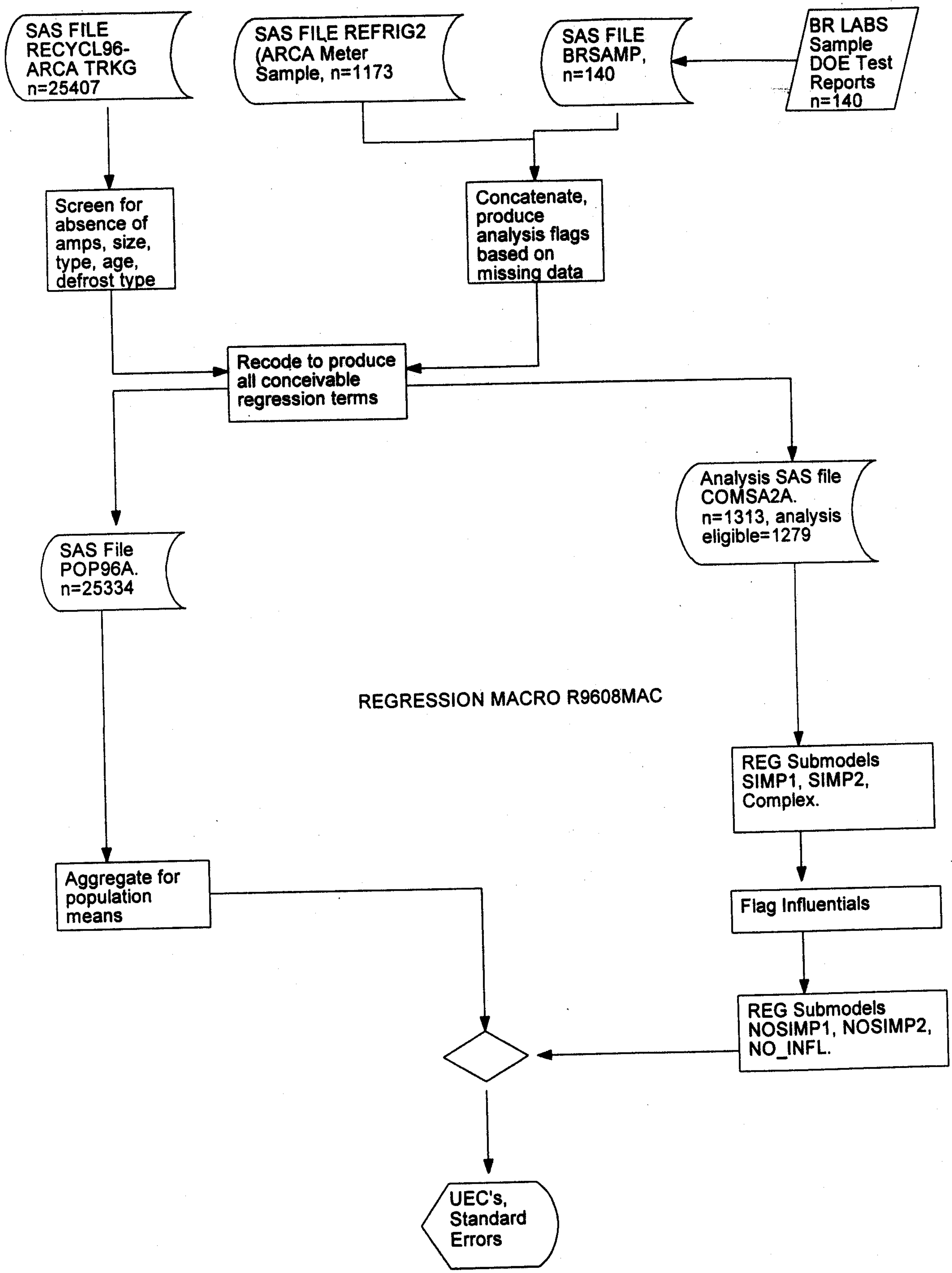
Table 3 is a simple tabulation reflecting the incidence of appliances in the 1996 population, the ARCA metering sample, and the BR metering sample, categorized according to a somewhat less detailed classification than the sampling plan used, in light of the small numbers of observations thus far occurring in parts of the sampling plan. Weights have been developed for this restratification (simplification of the orignal sampling plan's stratification), for use in the <u>illustrative</u> descriptive/comparative analysis contained in the next section. Note that the stratification and weighting issues are essentially inconsequential for the regression-based UEC estimation that is our primary goal. In that context, variance in the X-variables contained in the samples and mirrored by population averages is more critical. In Table 3, stratification variable STRAT2 is a concatenation of appliance type (F/R), defrost type for freezers (FF,MAN), configuration for frost free refrigerators (BF, SS, TF), a manual defrost/partial defrost meta-category (MDPD), and size groupings.

Table 3: Summary of Population, ARCA Sample, Achieved BR Sample, Restratified

FREEZ/ REFRIG	STRAT2	ARCA_WT	BR_WT	N_POP	N_ARCA	N_BR
0	F-FF___10-17	36.549	133.143	1864	51	14
0	F-FF___18+	65.481	160.727	1768	27	11
0	F-MAN__10-17	61.625	123.250	493	8	4
0	F-MAN__18+	175.200	125.143	876	5	7
1	R-BF___10-17	93.231	303.000	1212	13	4
1	R-BF___18+	137.917	275.833	1655	12	6
1	R-SS___14-20	10.184	221.500	886	87	4
1	R-SS___21+	39.729	146.500	2344	59	16
1	R-TF___10-13	21.714	76.000	152	7	2
1	R-TF___14-17	5.550	148.917	1787	322	12
1	R-TF___18-20	31.061	205.926	5560	179	27
1	R-TF___21+	19.944	51.286	359	18	7
1	R_MDPD_10-13	7.685	166.500	1998	260	12
1	R_MDPD_14+	46.105	438.000	4380	95	10
				=====	===	==
				25334	1143	136

Figure 1 is a flowchart documenting the highlights of data sources for this analysis.

Figure 1: Data Development and UEC Regression Procedures



III. ANALYSIS - DESCRIBING AND COMPARING THE POPULATION AND TWO SAMPLES.

This section provides simple descriptive and comparative information for the 1996 population (tracking system), the ARCA metered sample, and the BR achieved sample of 136 appliances. Sample data are weighted, using stratum-specific case weights given in the previous section. We compare appliance age, size, and amperage between these groupings, as well as an ARCA/BR comparison of metered consumption averages. The descriptive/comparative information contained here is not as definitive as the regression-based UEC estimation procedures described later in this document; however it is helpful to see the similarity of the samples and population.

Table 4, panel A, considers all appliances ("Overall"). The samples appear to be slightly younger (in weighted aggregate), than appliances in the population, but very similar with respect to average size and amperage. Each row of Table 4 presents a probability estimate ("alpha level") relating to the comparison of ARCA sample and BR sample weighted means. While the BR sample is slightly larger in terms of volume and has a slightly higher metered consumption, only the difference in age, favoring the BR sample, is significant (we will consider $p <= .10$ a liberal level for making such judgments).

Panel B simply disaggregates freezers and refrigerators for comparison. For refrigerators, both samples are more youthful than the population, and appear to mirror the population on other attributes. Refrigerators are strikingly <u>similar</u> between the two samples, on all variables including metered consumption. This is initial evidence against the argument that either selectivity or instrumentation biases in the ARCA sample threatened its validity for use in the evaluation. For freezers, evidence for these smaller samples suggests similarity of samples to population, and also indicates that the BR sample is younger and more energy-consuming than the ARCA sample (not significant differences).

Panel C of Table 4 goes a step further in detail, making comparisons within appliance types and configurations. Some age differences between samples achieve statistical significance, and there is a striking difference between samples' consumption for manual defrost freezers (F-MA).

On balance, however, the evidence suggests unbiased representation of the population, and seems to indicate the similarity of the ARCA and current BR samples. We note, based on Table 4, panels A and B, the "stand-alone" UEC's that would be obtained were the ARCA sample ignored and the BR sample used to develop a weighted mean (from the data thus far received and using the provisional weights developed here):

Overall	**2125**	**(standard error=89 kWh)**
Freezers	**2084**	**(190 kWh)**
Refrigerators	**2135**	**(101 kWh)**

Appendix E contains a fuller representation of the population/samples comparison, extending to detail on differences obtaining in size "substrata" within configuration types.

Table 4: Comparison of Population, ARCA Sample, BR Sample

A: OVERALL

OVERALL	Parametr	Populatn	ARCA Mean	ARCA Stderr	B.R. Mean	B.R. Stderr	p(diff)
1	COUNTS	25334.00	1143.00		136.00		
	AGE MEAN	23.26	22.03	0.25	20.92	0.59	0.0849
	SIZE	17.74	17.28	0.09	17.54	0.25	0.3110
	AMPS MEAN	5.43	5.54	0.04	5.55	0.12	0.9260
	ANNL KWH		2072.48	28.75	2124.67	89.03	0.5777

B: BY FREEZERS/REFRIGERATORS

Freezer/ Refrig	Parametr	Populatn	ARCA Mean	ARCA Stderr	B.R. Mean	B.R. Stderr	p(diff)
F	COUNTS	5001.00	91.00		36.00		
	AGE MEAN	24.25	26.08	0.73	22.36	1.21	0.0108
	SIZE	16.99	16.47	0.33	16.73	0.46	0.6541
	AMPS MEAN	4.77	4.83	0.11	5.07	0.17	0.2384
	ANNL KWH		1893.76	69.26	2084.14	189.93	0.3514
R	COUNTS	20333.00	1052.00		100.00		
	AGE MEAN	23.01	21.04	0.26	20.57	0.68	0.5180
	SIZE	17.93	17.47	0.09	17.74	0.29	0.3744
	AMPS MEAN	5.59	5.71	0.05	5.66	0.15	0.7738
	ANNL KWH		2116.44	31.76	2134.64	101.61	0.8646

Table 4: Comparison of Population, ARCA Sample, BR Sample (continued)

C: BY CONFIGURATION

Freezer/ Refrig	CONFIG3	Parametr	Populatn	ARCA Mean	ARCA Stderr	B.R. Mean	B.R. Stderr	p(diff)
F	F-FF	COUNTS	3632.00	78.00		25.00		
		AGE MEAN	25.13	26.92	0.86	24.00	1.51	0.1007
		SIZE	16.71	16.01	0.39	16.28	0.60	0.7060
		AMPS MEAN	4.47	4.53	0.11	4.70	0.18	0.4121
		ANNL KWH		1626.02	60.70	1580.31	114.46	0.7262
F	F-MA	COUNTS	1369.00	13.00		11.00		
		AGE MEAN	21.94	23.86	1.17	18.01	1.30	0.0031
		SIZE	17.70	17.71	0.47	17.92	0.46	0.7518
		AMPS MEAN	5.54	5.64	0.16	6.04	0.16	0.0905
		ANNL KWH		2604.09	103.91	3420.81	363.03	0.0521
R	R-BF	COUNTS	2867.00	25.00		10.00		
		AGE MEAN	25.64	22.39	1.46	23.23	2.67	0.7867
		SIZE	18.10	17.60	0.39	17.42	0.68	0.8237
		AMPS MEAN	6.45	6.33	0.17	6.29	0.16	0.8713
		ANNL KWH		2647.51	167.35	2645.66	222.56	0.9948
R	R-SS	COUNTS	3230.00	146.00		20.00		
		AGE MEAN	20.07	16.59	0.50	18.64	1.03	0.0843
		SIZE	21.52	21.36	0.17	21.89	0.54	0.3570
		AMPS MEAN	6.79	7.05	0.07	7.07	0.18	0.9173
		ANNL KWH		2950.34	92.85	2946.87	292.73	0.9911
R	R-TF	COUNTS	7858.00	526.00		48.00		
		AGE MEAN	19.88	16.60	0.27	17.66	0.73	0.1801
		SIZE	18.03	17.89	0.07	18.21	0.24	0.2173
		AMPS MEAN	6.01	6.25	0.05	6.08	0.16	0.3245
		ANNL KWH		2236.34	34.08	2233.52	131.50	0.9835

Table 4: Comparison of Population, ARCA Sample, BR Sample
(continued)

C: BY CONFIGURATION

Freezer/ Refrig	CONFIG3	Parametr	Populatn	ARCA Mean	ARCA Stderr	B.R. Mean	B.R. Stderr	p(diff)
R	R_MD	COUNTS	6378.00	355.00		22.00		
		AGE MEAN	27.26	28.15	0.41	23.93	1.51	0.0126
		SIZE	15.91	14.93	0.12	15.21	0.33	0.4429
		AMPS MEAN	4.07	4.09	0.06	4.16	0.22	0.7443
		ANNL KWH		1307.68	39.16	1371.76	94.33	0.5353

IV. ANALYSIS - REGRESSION BASED ESTIMATES OF ANNUAL CONSUMPTION.

We begin with the model selected, and later discuss the development of this model, alternatives considered, and diagnostics. A general point to be emphasized is that this two sample approach is providing the UEC estimation with a large and sufficiently varied set of data points to allow estimation of a regression equation that serves as a "simulation engine" to which population data can be applied in achieving various point estimates, including but certainly not limited to overall UEC's or UEC's by appliance type. The two sample character of the analysis allows checking for and adjusting for any bias that may have been at work in the ARCA metering sample. As it turns out, this bias is not an issue.

A. Final Model.

The model selected was developed from a limited hierarchy involving several carefully varied alternative models. In this section, we present the selected model, and discuss its content and some diagnostic efforts. In later sections we present its resultant UEC estimates, and the progression of the model development.

Table 5 is SAS output describing the final UEC model run, model C1_NOINFL. We will begin by describing the context of the model, the regressors identified for this model and the the rationale for their inclusion, and interpreting certain of the key coefficients.

The model was estimated using OLS over a combination of 1143 appliances from the ARCA study sample and 136 appliances from the BR study sample. This includes all the appliances with nonmissing data on metered kWh, age, volume, amperage, defrost style, and configuration, with the exception of two records excluded after a prior run of the same model, in which the DFFITS values (case-specific influence diagnostics) for these records exceeded a liberal criterion of 1.2814. These two records, each from the ARCA metering sample, also had very high metered consumption, exceeding 8000 kWh/year (tested appliances with smaller consumption and peculiar combinations of regressors could easily have been highly influential as well). It will be noted that the starting point for these regressions (described later) was somewhat influenced by previous experience with these data in connection with a prior estimation effort carried out for Edison staff, before the BR sample was near completion.

In order to identify the variety of appliance types, defrost type, and configuration for the regression, a number of dummy variables are included:

FREEZER is a binary reflecting that the appliance is a freezer.

D_FF is a binary reflecting frost free defrost as opposed to manual or partial defrost.

D_SD, D_SS, D_TF, D_UF are binaries reflecting single door refrigerators, side-by-side refrigerators, top mount freezer refrigerators, and upright freezers.

D_FFBF, D_FFSS, D_MDSD, D_PDTF continue by representing configuration-by-defrost combinations frost free-bottom freezer, frost free side-by-side, manual defrost single-door, and partial defrost top freezer.

The above binaries fairly nearly exhaust the detail that the regression model can support over the appliance types observed in the samples.

The variable AGERT is equal to the square root of the appliance age, and is used to express the possibly declining significance of age in appliance degradation (or "cohort" effects relating to manufacture date). Variable AGE (the actual number of years) is used in other specifications, with trivial impact on results. AMPS gives the rated (label) amperage of the appliance. SIZE is the volume (cubic feet) recorded in the tracking data by ARCA. Age and size variables are obvious candidates for a regression purporting to explain appliance consumption. Note that size, defrost type, configuration, age, and amperage are all variables from ARCA tracking, so that there is a consistency in meaning between samples and also in the later application of population data from the 1996 tracking data to the calibrated regression.

Table 5: Model C1 SAS Output

Model: NO_INFL
Dependent Variable: ANNKWH2

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	20	630555578.17	31527778.908	61.080	0.0001
Error	1256	648311423.62	516171.51562		
C Total	1276	1278867001.8			

Root MSE	718.45077	R-square	0.4931
Dep Mean	1946.51830	Adj R-sq	0.4850
C.V.	36.90953		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|	Type I SS	Variable Label
INTERCEP	1	1138.067612	749.98046277	1.517	0.1294	4838468073	Intercept
FREEZER	1	-932.026917	807.56057632	-1.154	0.2487	1181152	FREEZER DUMMY
D_FF	1	319.673831	289.87492419	1.103	0.2703	470758746	DUMMY - DEFROST - FROST FREE
D_SD	1	-541.441203	728.83635585	-0.743	0.4577	51518065	DUMMY - SINGLE DOOR
D_SS	1	-717.994230	794.83958835	-0.903	0.3665	45844682	DUMMY - SIDE BY SIDE
D_TF	1	-810.609774	714.91706705	-1.134	0.2571	9172140	DUMMY - TOP FREEZER
D_UF	1	406.706157	342.47510046	1.188	0.2352	821822	DUMMY - UPRIGHT FREEZER
D_FFBF	1	-672.632544	329.40115725	-2.042	0.0414	1048450	DUMMY,CONFIG= FR FREE BOTTM FRZR
D_FFSS	1	843.248396	429.97283927	1.961	0.0501	8161918	DUMMY,CONFIG= FR FREE SIDE X SIDE
D_MDSD	1	-361.554510	223.91511650	-1.615	0.1066	2093478	DUMMY,CONFIG= MANL DEF SINGL DOOR
D_PDTF	1	457.281809	146.88068104	3.113	0.0019	8318585	DUMMY,CONFIG= PARTL DEF TOP FRZR
AGERT	1	-36.990504	43.66643541	-0.847	0.3971	2353805	age root
AMPS	1	109.170909	23.48310820	4.649	0.0001	18707002	
SIZE	1	30.390039	10.79428054	2.815	0.0049	3672003	

Table 5: Model C1 SAS Output
(Continued)

D_BFD	1	24.882448	135.36325912	0.184	0.8542	28331	bott freez x amps
D_SSD	1	-65.221623	69.52828610	-0.938	0.3484	701129	side-side x amps
FREEZERD	1	44.602989	71.60843239	0.623	0.5335	443814	freezer x amps
D_FFC	1	137.057372	54.23694132	2.527	0.0116	3117234	frost free x agert
NEWSAMP	1	-738.622520	392.80416078	-1.880	0.0603	50095	Dummy for new sample member
NEWFREEZ	1	212.427503	177.21661258	1.199	0.2309	999540	interact newsamp*freezer
NEWAGERT	1	154.099560	88.53952531	1.740	0.0820	1563588	interact newsamp*agert

Variables D_BFD and D_SSD are interaction terms meant to allow for the possiblity that label amperage's impact upon consumption may be somewhat different for larger and more complex appliances (bottom freezer and side-by-side refrigerators), and FREEZERD allows for a possibly different relationship among freezers.

The variables thus far described are considered "theoretically non-negotiable" in accounting for consumption, meaning that we feel their inclusion in the model is important enough that we would not "trim" these regressors from the model and distort its meaning based on, say, significance testing.

We did entertain some interactions on a contingent basis, however. D_FFC is the interaction of AGERT (square root of age) and D_FF (a dummy reflecting frost free appliances). Coupled with the additive specification of AGERT, the suggestion from the coefficients is that age is consequential for frost free appliances only (a possible artifact of the relationship between age and frost free defrost).

Finally, we include three variables allowing for differences between the newly collected BR sample and the ARCA metering sample. NEWSAMP is an additive term reflecting BR sample membership. Its large negative coefficient here is not to be understood as the simple difference in adjusted means between the two samples, given the inclusion of NEWFREEZ and NEWAGERT, which indicate the larger effect of freezers (vs. refrigerators) and appliance age upon consumption in the new sample. As will be explained later, simple additive specification of BR vs. ARCA sample effects using NEWSAMP alone leads to a trivial coefficient on this variable, consonant with the results described in an earlier section comparing the two samples. We include these three regressors (NEWSAMP, NEWAGERT, and NEWFREEZ) in order to be conservative with respect to sample differences (i.e., to account any main or interactive effects that might be otherwise biasing the regression). These variables, as explained later, were selected from among other similar interaction terms because of the individual significance of NEWAGERT and NEWFREEZ in a previous regression which tested several interaction terms, and the marginal joint significance of the pair's inclusion in the model.

Here, we will discuss certain diagnostic work with respect to this final model, and reserve other diagnostic issues for a more general discussion of the modeling effort.. Regarding the final model, we will discuss the issues of influential data points, specification (revisited it during a description of model development), and multicollinearity.

Influential data points. In the development of the model, we set a DFFITS criterion of 1.2814 for the exclusion of data points. This cutoff value balances concern for the inclusion of all valid observations with our interest in developing a simulation engine from these samples which includes a comprehensive and robust representation of how old refrigerators' consumption is affected by a handful of substantive variables. The cutoff value is approximately five times the value suggested by Belsey, Kuh, and Welsch (1980), but considerably less than the general cutoff of 2.0 used by other practitioners. At any rate, throughout the model development, including estimation of final model C1_NONINFL, one of the precursor tasks to any model estimation was to perform the same model run with all valid (nonmissing data) records, and rerun ("_NONINFL") the model with influentials removed by setting the dependent variable ANNKWH2 to missing. Through the progression, two particular observations were typically removed. These particular appliances came from the earlier ARCA sample, and have the following attributes (Table 6):

Table 6: Influential Observations.

Defrost	Config.	Size	Amps	Age	Annlkwh
FF	Top freezer	20	5.0	12	8714
FF	Side-side	24	8.5	10	9145

It will be noted that the overall UEC change from the all-inclusive precursor model to final model C1 amounts to a drop of one kWh, reflecting that (a) the model is robust, and (b) the input values from the tracking system population are highly determinative of the UEC estimate, once a simulation engine has been built from samples containing appropriate variability in the regressors.

Specification. Earlier, we described the rationale for inclusion of regressors in the final model above. To follow up, we correlated residuals from the final model with a comprehensive set of omitted variables in the analysis data set (mainly omitted interaction terms, linear or nonlinear treatments of included interval level variables, etc.). Very few omitted terms correlated with the residuals at all. A very few variables specifying higher-order interactions involving configuration-age category-amperage interactions correlated in the .035-.05 range. We do not take these results to be indicative of mis-specification. The matrix revealing these findings is contained in Appendix F, combined with some correlational information relevant to heteroscedasticity concerns.

Heteroscedasticity. When the expected magnitude of regression residuals relates to model regressors, the reliability of standard errors from the regression may be suspected, and in some cases GLS solutions may be appropriately considered in lieu of the OLS estimation (obtaining in this analysis). These efforts will not result in superior parameter estimates, however, and may in fact lead to specification errors. We correlated the absolute value of model residuals (from final model C1) with regressors, and found that there was in fact some evidence of heteroscedastic error. Indicators of heteroscedastic error are found in Appendix F, where the absolute value of residual error (ABSR2) is correlated with model variable. The larger and significant correlations are provided below. These data do suggest the possibility of downwardly biased standard errors for variables relating, directly or indirectly, to the size of these appliances, and that further effort could include an extension into GLS.

Table 7: Correlating Absolute Residuals with Model Variables

Variable	r
D_FF	0.19
D_SD	-0.16
D_MDSD	-0.17
AGERT	-0.16
AMPS	0.18
SIZE	0.19
D_FFC	0.16
ANNKWH2	0.39

Multicollinearity. The model estimation exhibits some multicollinearity; however all of it is due to the "logical multicollinearity" which follows of necessity when a number of terms are being used to capture a few categorical variables and their interactions. The degradation owing to multicollinearity is certainly more than compensated for by the information content in the necessary inclusion of this categorical information (as an example, one is unlikely to conduct a commercial impact evaluation without including building type's inclusion; if the computational expedient of creating building type dummies is carried out, some considerable but highly worthwhile multicollinearity ensues). Table 8 lists and comments on the contents of the smallest ten unique variance/covariance components among the regressors, with correspondingly high condition indices, and the regressors for which any given unique component contributes thirty or more percent of estimate variance (two or more such contributions by one component suggests that multicollinearity may be affecting the stability of the estimates, which is reflected in estimate standard errors).

Table 8: Smallest 10 Unique Variance Components
(Top 10 Collinearity Indices)

#	Index	Regressors, (Variance Proportions), Comments
12	15.1	D_UF (0.5) FREEZERD(0.3) NEWFREEZ (0.1) freezer/upright/new freezer sub-categorization impacts not completely independent
13	16.8	D_MDSD(0.8) unique impacts of manual defrost single door
14	18.7	D_UF(0.2) AMPS(0.6) upright freezers and amperage
15	21.2	D_FFSS (0.2) FREEZERD(0.2) slight interplay between estimation of frost free side-side and freezer main effect
16	25.2	D_FFSS(.4) D_PDTF (.3) FREEZERD (.2) NEWSAMP (.1) NEWAGERT(0.1) interplay between frost free side, partial defrost refrigerators, freezer, new sample main effect, and unique impacts of age in new sample
17	26.0	NEWSAMP (.7) NEWAGERT (.7) interplay between new sample binary and its correlated interaction with age
18	29.1	SIZE (.8) unique impacts of size upon consumption
19	29.8	D_FFSS (.4) AMPS (.2) D_SSD (.9) interplay in estimating amperage impacts and the effects of particularly large refrigerator configurations
20	49.3	D_FF 8 AGERT 8 D_FFC 8 frost free freezers/refrigerators, age, and age by frost free interplay among arithmetically related terms
21	138.1	INTERCEPT 9 FREEZER 7 D_SD 9 D_SS 8 D_TF 9 D_BFD 9 intercept adjustments due to certain major configuration types

B. UEC estimates based on model C1.

To obtain UEC estimates from the model and the 1996 tracking system population, we applied mean population values on all model regressors to the calibrated final equation (model C1_NOINFL) – means obtaining for refrigerators, for freezers, and for all appliances. In each case, the standard error for the UEC thus obtained is equal to

$$se\_UEC = (MCM')^{1/2} ,$$

where M is a vector of means and C is the variance-covariance matrix of regression estimates. To allow for the modest impacts of the BR sample vs. those of theARCA sample, the population mean corresponding to BRSAMP is set to 1.0, and interaction terms are set to equal the means on the relevant quantitative variable which is interacted with BRSAMP in the modeling.

The UEC values obtained are given in Table 9. The second panel gives the net increment or decrement in kWh attributable to the new sample (new sample membership and allowance for differential slopes in the new sample), and the standard error of this increment.

Table 9: UEC Values

Appliance	UEC (kWh)	Std. Error	Relative Precision: 90%	80%
Freezers	2058.301	125.882	10.06%	7.83%
Refrigerators	2147.594	80.036	6.13	4.77
Overall	2130.077	69.658	5.38	4.19
New sample impacts:				
Freezers	222.132	161.387		
Refrigerators	-9.200	85.365		
Overall	36.456	77.848		

C. Model development—alternatives.

As indicated earlier, the progression toward model C1 began from a base which was already informed somewhat by previous work with the ARCA sample and a much smaller subset of the BR sample (those complete as of the end of March 1998). Here, we briefly describe starting model A; alternatives B, B1, B2 involving handling of appliance age; alternatives C1, C2, C3 involving the specification of BR sample effects; D1, D2 in which highly simplified alternatives are tested for information purposes only. Each of these regressions is displayed in Appendix G. In this section, we limit ourselves to discussion of the content of the models, the basis for decisions based on it, and a table giving UEC estimates based on the various calibrated results.

Each model run was actually implemented as a set of six "sub-runs":

SIMP1: all model variables except NEWSAMP (membership in BR sample), and interactions between NEWSAMP and other model variables.
SIMP2: all model variables, including NEWSAMP, eliminating NEWSAMP interactions.
COMPLEX: all model variables including NEWSAMP and its interactions.

NOSIMP1: SIMP1 rerun, dropping influential cases (DFFITS exceeding 1.28) from model COMPLEX.
NOSIMP2: SIMP2 rerun, dropping influential cases identified in model COMPLEX.
NO_INFL: COMPLEX rerun, dropping influential cases identified in that model run.

The intent of the "sub runs" was to establish hierarchically the role of NEWSAMP and its affiliated interactions, and to establish differences emerging after eliminating influential records.

Model A. Run over 1279 records, Model A included configuration and defrost terms similar to those described for model C1. It also included dummy variables reflecting particular size categories, AGERT, amperage, an interval expression of size, interactions between configuration and amperage, and interactions between configuration/defrost type and age. The COMPLEX and NO_INFL sub runs included additive BR sample effect NEWSAMP as well as four potential interactions involving (1) frost free appliance, (2) freezer, (3) amperage, and (4) age (square root of age). Two observations with very high consumption were omitted in the second round of sub runs, with COMPLEX achieving an Rsquare of 0.4972. Inspection of results suggested that variable FREEZERC (interaction of freezer appliance and AGERT) should be dropped due to miniscule statistical significance, but that the other interaction involving AGERT – D__FFC or interaction of frost free defrost and AGERT should be retained.

Model B. This model deleted variable FREEZERC, yielding an Rsquare value of 0.5011. This slight gain is due to the fact that three rather than two observations were dropped based on influence values, leading to a slightly better fit. In this case, three observations were dropped by the DFFITS-filtering routine. The model included three age group dummies to allow for possible "era" effects in addition to a linear age impact (or at least a monotonic one in the case of AGERT). Attention turned to whether these dummies should be retained.

Model B1. This model deleted categorical age groups as expressed by dummies, yielding an Rsquare of 0.5000. The F value associated with this relaxation of the handling of age indicated that little explanatory power was sacrificed (F3,1275=0.85). We then considered whether AGERT (square root of age) was in fact better replaced by a simple linear expression of age (in years).

Model B2. With very little change in coefficients on other terms, the move to raw AGE (and interactions with frost free status and with BR sample membership) diminished the Rsquare slightly (to 0.4981), suggesting that essentially either specification of age (nonlinear, declining impact or linear in years) would suffice.

We opted to return to model B1, and assess whether the entire set of BR sample interaction terms should be retained. Scanning coefficients (rather than taking the somewhat more rigorous approach of performing a Chow test on the entire set), we observed that variables NEWAMPS and NEWFF were of trivial significance, and perhaps, based on collinearity, degrading the solution.

Model C1. With Rsquare dropping from 0.5000 (Model B1) to 0.4931 (Model C1), we observed that sample membership makes little additive difference at all (NOSIMP2 vs NOSIMP1; additive term NEWSAMP's coefficient is -22.022 kWh). Adding in the two remaining interaction terms (not trimmed from model B1 – NEWFREEZ and NEWAGERT) makes some difference, in that

(a) an F test for the addition of additive and interactive terms (NEWSAMP, NEWFREEZ, NEWAGERT) to a base model including no sample membership terms (sub model NOSIMP1) is 1.811 (p= .14).
(b) an F test for the addition of the interaction terms to a base including NEWSAMP (i.e. moving from NOSIMP2 to COMPLEX) is equal to 2.588 (p= 0.08).

We settled on model C1 as a reasonable alternative to any further testing, noting that further attention to specification changes on this data set, without an alternative validation data set, would be essentially a fishing expedition. Given the concern about the ARCA sample's potential role in biasing metering study results, we also felt it would be best to end with a model including a direct attack on the issue, even though coefficients for the sample membership-related variables were of marginal joint significance. We did run models C2 (drop NEWSAMP interactions; identical to running submodel C1.NOSIMP2) and C3 (drop NEWSAMP and interactions; identical to running submodel C1.NOSIMP1), simply to obtain alternative estimates of the UEC with models that ignore sample differences.

Finally, we ran simplified models D1 and D2 as checks to verify that the details in models A—C1 were in fact of some explanatory consequence. Model D1 included only basic dummies reflecting configuration and defrost, AGERT, AMPS, and the NEWSAMP dummy, while Model D2 discarded NEWSAMP from the specification of D1. These model results are also included in Appendix G.

D. UEC Estimates Based on Alternative Models.

In this section, we present the UEC estimates that were obtained from each of the models described above, including final model C1.

Table 10 provides UEC estimates derived from the various models, including the final model. It will be noted that the models generally provide estimates above 2000, for both freezers and refrigerators. This accords with sample mean estimates introduced earlier, and suggests the robustness of the approach under changes in specificaton.

All joint BRSAMP effects, whether simple additive impacts in the case of model C3 or model D1, or additive/interactive, are statistically insignificant. However there are appreciable sample effects upon freezers. The pattern of refrigerator UEC estimate standard errors shows that removal of terms relating to BR sample status increases the precision of estimates (which does not mean, however, that the terms and the issues of sample differences should not be directly addressed as they are in regression C1).

Table 10: UEC Values, Various Models
kWh/ (standard error)

Model	UEC (kWh) Freezer	UEC (kWh) Refrig	UEC (kWh) Overall	NEWSAMP (plus interactions) kWh impacts Freezer	NEWSAMP (plus interactions) kWh impacts Refrig	NEWSAMP (plus interactions) kWh impacts Overall
a	2065.9 (124.6)	2083.9 (80.0)	2080.4 (69.5)	230.5 (166.8)	-71.8 (85.0)	-12.2 (77.4)
b	2061.1 (124.3)	2088.7 (79.5)	2083.3 (69.3)	204.6 (160.4)	-66.7 (84.5)	-13.2 (77.4)
b1	2060.9 (124.2)	2082.9 (79.3)	2078.7 (69.1)	214.0 (160.2)	-73.5 (84.5)	-16.8 (77.4)
b2	2057.1 (124.5)	2077.5 (79.7)	2073.5 (69.5)	217.2 (161.2	-67.9 (84.7)	-11.6 (77.3)
c1	2058.3 (125.9)	2147.6 (80.0)	2130.1 (69.6)	222.1 (161.4	-9.2 (85.4)	36.5 (77.8)
c2	1937.9 (100.6)	2133.7 (34.9)	2095.2 (64.0)	-22.0 (70.8)	-22.0 (70.8)	-22.0 (70.8)
c3	1948.2 (95.0)	2152.0 (34.9)	2111.9 (35.0)	(-- no NEWSAMP terms --)		
d1	1922.6 (106.0)	2129.7 (71.2)	2102.1 (67.0)	-17.3 (73.8)	-17.3 (73.8)	-17.3 (73.8)
d2	1931.0 (99.8)	2144.2 (35.2)	2102.1 (35.6)	(-- no NEWSAMP terms --)		

REFERENCES

1. Belsey, D.A., Kuh, E., and Welsch, R.E., Regression Diagnostics. New York, John Wiley and Sons, Inc., 1980.

Southern California Edison, 1998. Refrigerator Recycling Program Metering Project. Conducted by BR Laboratories, Huntington Beach, CA.