Trying Too Hard to be Cool: An Exploration of Refrigerator Degradation

Miriam Goldberg, KEMA Inc.
Ken Agnew, KEMA Inc.
Shahana Samiullah, Southern California Edison

Abstract

This study explored the change in refrigerator Unit Energy Consumption (UEC) with age, using metering data from a sample of units collected by California's Residential Appliance Recycling Program (RARP). The objective of the analysis was to determine the extent to which the UEC of a given refrigerator or freezer production model recycled under the RARP increases (or decreases) with age, and how that age effect is related to other characteristics of the refrigerator.

This question has implications for both design and impact evaluation for recycling programs. If the UEC increases appreciably as refrigerators age, the savings from these programs, and corresponding cost-effectiveness, is greater than might otherwise be projected. In particular, savings from recycling units manufactured since the advent of new efficiency standards in the early 1990s are projected to be relatively low. These newer units are currently ineligible for the California RARP. However, if UECs increase sufficiently with age, even these post-standards units may be cost-effective to collect as they age.

The analysis indicates that degradation, in terms of increased UEC, does take place in refrigerators and freezers over their lifespan. The analysis further indicates that degradation can be quantified as a function of unit characteristics. In particular, the effect of age varies with unit configuration and size.

Limitations of the data preclude definitive quantitative estimates for the current program. However, the work demonstrates the ability of this approach to provide improved projections of potential savings for a program population, and determine the best cut-off date for program eligibility.

Introduction

California's statewide Residential Appliance Recycling Program (RARP) provides environmentally safe removal and disposal of "excess" residential refrigerators and freezers. The energy savings from units turned in to this program is a function of the average full-year Unit Energy Consumption (UEC). A question with implications for both design and impact evaluations of this type of program is how the UEC changes over the life of a refrigerator. If the UEC increases appreciably as refrigerators age, the savings from these programs, and corresponding cost-effectiveness, is greater than might otherwise be projected.

Like many such programs around the country, the California RARP recognizes energy savings from two sources:

1. eliminating the use of a spare refrigerator or freezer that the participant would otherwise have continued to operate and
2. avoiding the transfer of a participant's old, unwanted unit to a new home.

In the first case, the savings credited to the program is the annual energy use of the eliminated unit. In the second case, the savings credited to the program is the difference between the consumption of the old unit collected by the program and that of the newer unit the would-be recipient acquires in its place, if any.

The most recent evaluation of the California RARP (KEMA Inc., 2004) found that the great majority of units collected by the program fell into the latter category of "avoided transfers." Savings from these units are large enough to make the program successful largely because the units turned in were mostly manufactured prior to the new energy efficiency standards of 1990 (California) to 1993 (federal), while the alternative unit was typically manufactured under the new standards. For a post-1990 unit turned in to the program, the estimated savings would be small.

In light of the low projected savings from newer units, the California RARP rules were revised last year to exclude units manufactured later than 1989. However, if a unit's energy consumption increases sufficiently over time, units manufactured since the new standards went into effect and recycled 10 or more years later could still provide enough savings to justify the program cost.

This study explored the change in refrigerator UEC with age, using metering data from a sample of units collected by the RARP. The objective of the analysis was to determine the extent to which the UEC of a given refrigerator or freezer production model that is recycled under the RARP increases (or decreases) with age, and how that age effect is related to other physical conditions of the refrigerator. To address this issue, we compared the manufacturer-reported UEC of a given production model ("new") to that of a unit of the same production model ("old") that was recycled by the RARP. Both the manufacturer's rating of a new unit and the measured UEC of an old unit collected by the program are based on the same Department of Energy (DOE) laboratory test protocol.[1] We also tested analytic models that would express degradation, in terms of increased UEC, as a function of age and other unit characteristics to attempt to quantify the effect of age on UEC for units turned in to the program.

Data

The central data for this analysis came from two samples of units, each including refrigerators and freezers, collected by the California RARP. In 1998, when the program was a Southern California Edison program, 140 units were collected, and in 2003, as a statewide program, 100 units were collected. The units were metered according to DOE protocol and were originally used to estimate UEC for the 2002 impact assessment. For this analysis, the combined sample of metered units represented RARP recycled units when "old."[2]

Obviously, we were not able to observe each of these units "as new." However, one advantage of DOE protocol laboratory metering is that it is reasonably consistent over time. When manufacturers produce ratings for their products, the ratings are based on estimates using DOE protocol metering as well. Thus, we can use the manufacturer ratings for a particular production model as an estimate of usage "as new" for each production model in the RARP data. Combining these two sources of data we have UECs for units of the same production model "as new" and when "old."

This analysis relied on our ability to match metered production model numbers from the "old" units to databases containing manufacturer estimates of "as new" usage. "Old" unit model numbers were available for a total of 240 units. We used two different sources for the manufacturer rating of "new" usage. The primary source was the California Energy Commission (CEC) regulatory data made available in electronic form by the Weatherization Assistance Program Technical Assistance Center (WAPTAC). These data include refrigerator production model numbers and usage for refrigerators sold in California. The Association of Household Appliance Manufacturers (AHAM) also maintains a

---

[1] A separate question affecting program design and evaluation is how well the laboratory UEC measurement compares with actual usage in a home. We reviewed this question based on existing literature in the 2002 evaluation. The question is being assessed empirically as part of the current evaluation effort in a new metering study comparing *in-situ* with laboratory UEC measurements for the same units.

[2] Code of Federal Regulations (CFR), Section 430.23 (a), (2001).

database in support of its volume certification program. We used the AHAM data to extend and fill in the WAPTAC/CEC data. The two sources include data for units manufactured from 1975 through 1977 and 1979 through 1992. Thus, data were not available for units made before 1975 or after 1992, nor for those made in 1978. In actuality, the coverage is wider than just those years for which we have the manufacturers' databases because we matched by model number, and models are frequently produced for a number of years.

There are several limitations to the manufacturers' databases.

1. **Manufacturers' databases are incomplete.** Not all RARP sample brands are represented in the available manufacturers' data. Manufacturers provide usage data to AHAM on a voluntary basis. AHAM indicates that one of the major manufacturers dropped out of the organization after 1992. Other smaller manufacturers are also not present. The CEC/WAPTAC data would not necessarily include units brought into California from other states. Since we use primarily WAPTAC/CEC and pre-1992 AHAM data, missing manufacturers were not a major issue.
2. **Fuzziness of manufacturers' rated UEC.** The manufacturer's rated UEC for a given model number may not be based on the usage of the exact same production model. The rated UEC is not necessarily based on metered units of that particular model number, but may be an extrapolated value based on metering of similar production models. Also, rated UECs are not the same across years for the same production model. This difference in rating for the same production model could reflect actual changes in the production model's UEC or changes in the estimate.
3. **Model number length.** The manufacturer's data does not provide as many digits as many of the recorded model numbers for the metered units. In general, it appears the additional digits relate to characteristics such as color that do not affect the UEC, but it is not always possible to be sure.

Matching

Model numbers of RARP sample units had to be matched to the manufacturer databases one at a time. The model numbers for the RARP sample units were originally recorded in hard copy by the metering laboratory. The numbers were then manually entered to attach them to the electronic tracking information. Model numbers are long and complicated, containing a mix of letters, numbers and other symbols. When model numbers are transcribed from nameplates, it is easy for an S to become a 5, or an O a 0. This potential for transcription errors added a wildcard into a matching process that was already challenging given that the manufacturers' data did not provide as many digits as many of the recorded model numbers for the metered units. Because of these and other irregularities, the matching was done manually and involved a degree of judgment.

We were successful in matching 136 of the 240 metered units to the databases containing "as new" UECs. Of the 136 matches, 96 model numbers were identical to the model number digits provided in the manufacturer information databases. We refer to these as "complete matches." The other 40 differed by one or two digits and appeared to be a reasonable match based on configuration and defrost information on the "old" metered unit. We refer to these as "close matches." We considered these two groupings as we looked at statistics and tested regression models.

Age

Of fundamental importance in this analysis is the age of the units under consideration. The program-tracking database includes an estimate of each unit's age made by the recycling company hauler. This is the age data we have used when performing the impact evaluation and retention studies of this program. It has always been clear that the assigned ages are an estimate. The vast majority of units have assigned ages that are multiples of five.

The matching process forced a comparison between the age data from the tracking database and the ages indicated by the manufacturers' databases. Since a production model may be manufactured over a span of years, each RARP unit could potentially match with a span of years. Even given this flexibility, the disparity between age from the tracking database and age as indicated by the manufacture year was dramatic. As important as age was to this analysis, we felt we had to use the age data that was most grounded in reality. The manufacturers' database was clearly a reliable source of when a production model was available for purchase. On the other hand, there was no way to know (even using the tracking age data) when the unit was purchased during the span of years the model was produced.

A decision was made to use the manufactures' data to determine two different age variables that would bracket the span of possible ages. We refer to these as the maximum and minimum production age as they represent the beginning and end of the production cycle in the manufacturers' ratings available to us. For the models produced over multiple years, the UEC may or may not change. One production model included in this analysis had a UEC that dropped 18 percent over the span of its production. Bracketing the span of production years allowed us to also bracket the UECs if they changed. Using this approach, we were able to maintain the consistency of the manufacturers' data while remaining flexible to the range of possible ages. A production model produced for only one year was given the same minimum and maximum production age and a single UEC.

Figure 1 plots "tracking age" (the age reported in the program-tracking database) versus the maximum production age (based on model number). The regression line in the figure and the scatter around that line indicate a positive but weak relationship between the estimated age in the tracking database and the maximum production age. Furthermore, there is little or no difference between this plot against maximum production age and the same plot (not shown) using minimum production age. The manufacturers' production spans may not be complete, but they are reality based. The considerable scatter evident here indicates that the tracking database estimate of age is not a reliable indicator of unit age.

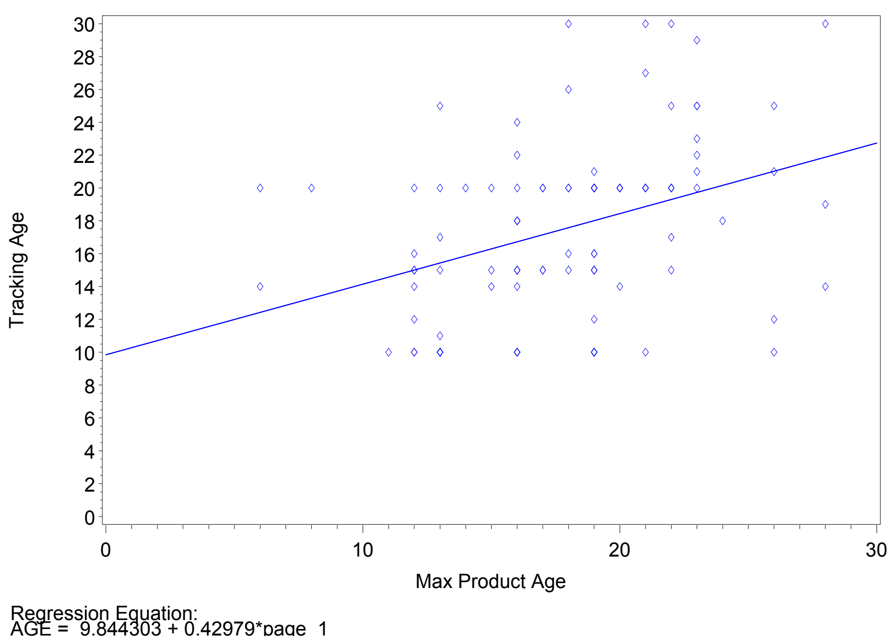

Figure 1. Tracking Age Versus Maximum Production Age, Complete Matches

The plot also shows that all but three of the units available for this analysis are older than 10 years, using the production age measure. Thus, these data will offer little information on degradation for units from new to age 10.

Calculating Degradation

For the 136 matches, we compared the “as new” and “old” UEC in terms of both the difference in UEC and the ratio of old to new UEC. The distribution provides a clear indication that degradation does occur. Of the 136 units under consideration, only 7 percent or fewer had “old” UECs lower than the “as new” UEC. The median difference between “as new” UEC and “old” UEC was either 557 kWh or 591 kWh depending on whether the maximum or minimum production age was used. The corresponding median ratios were 1.46 and 1.47. Thus, the “old” units overwhelmingly used more energy than the estimated UEC when new.

Figure 2 shows the distribution of the UEC differences. The bins are 200 kWh wide (the label indicates the bin midpoint). These are the differences when using maximum production age.

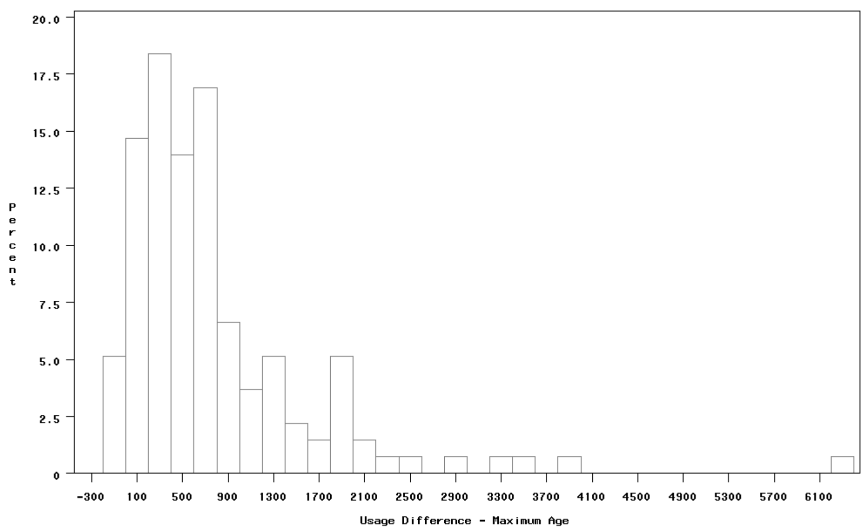

Figure 2. Distribution of UEC Differences Using Maximum Production Age (Earliest Possible Match)

Figure 3 shows the UEC difference in proportional terms—old metered data as a proportion of the new usage. The bins are 0.1 wide.

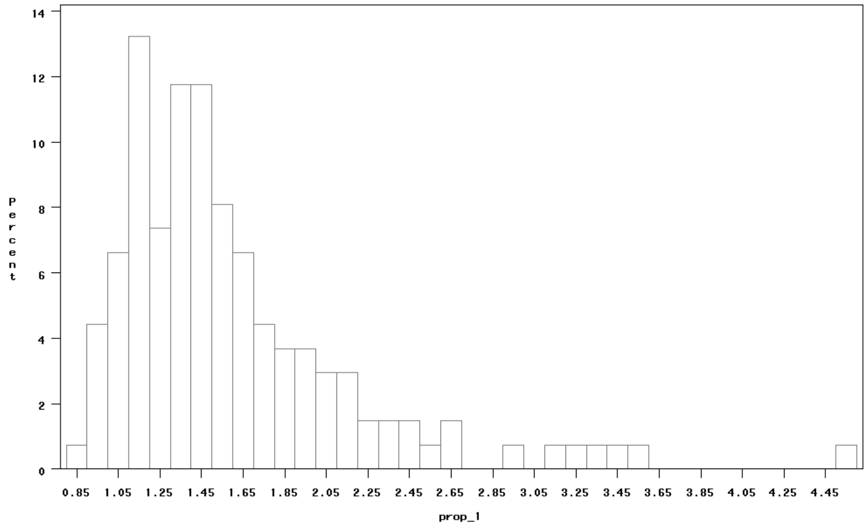

Figure 3. Distribution of Proportions Using Maximum Production Age (Earliest Possible Match)

Regression Analysis

A goal of the degradation analysis was to develop a regression model that expressed unit degradation as a function of age and other unit characteristics. A preliminary examination of plots generally provides a good indication of the likelihood of success with regression analysis. Figure 4 plots UEC difference versus the maximum production age for the units with complete model number matches. The different symbols (and colors) represent the two different samples of metered units.

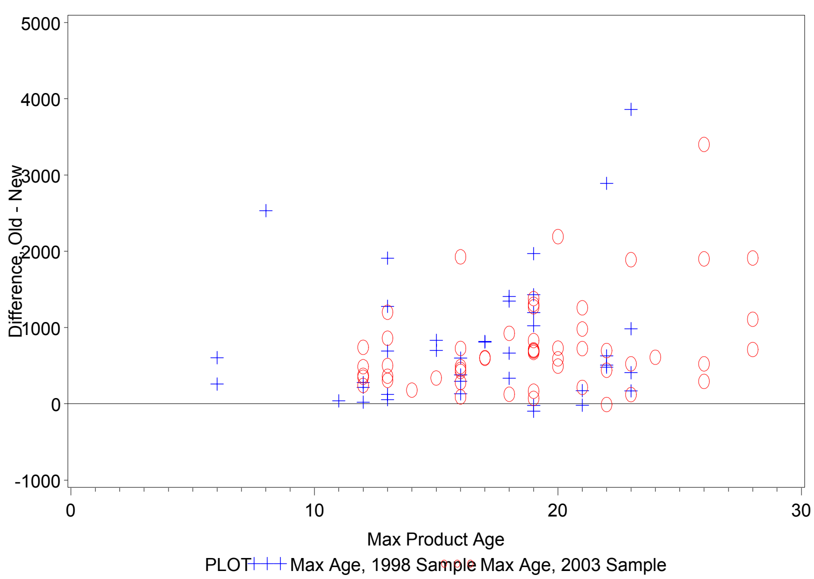

Figure 4. Usage Difference by Maximum Age, Complete Matches

Figure 4 appears to show a weak age-related trend. That is, the old-new difference in UEC tends to increase with age. The trend appears somewhat stronger in the 2003 sample. The trend is more pronounced in this plot of only the complete matches than if the close matches were included. Of course, Figure 4 includes the effects of configuration, size and even whether the unit was a refrigerator or freezer. Using regression analysis allowed us to control for unit characteristics.

Basic Regression Model Forms

We tested the relationship between old and new UEC using three different functional forms. The basic functional form used for the regression analysis was:

$$\#1)\ ln(\mathrm{OLDUEC}_j) - ln(\mathrm{NEWUEC}_j) = a + b_1 t_j + b_2 c + e_j$$

where t is age and c is some set of unit characteristics. The difference of the logged "old" and "as new" usage produces a dependent variable that approximates the percentage increase of UEC from "as new" to "old." The approximation is good when proportions are closer to one but are increasingly scaled down as the proportions increase beyond one.

Two variants were also tested:

$$\#2)\ ln(\mathrm{OLDUEC}) = a + b_1 t + b_2 c + b_3 ln(\mathrm{NEWUEC}) \text{ and}$$
$$\#3)\ \mathrm{OLDUEC} - \mathrm{NEWUEC} = a + b_1 t + b_2 c.$$

The first of these additional functional forms was essentially the same as model #1 above except that by moving the "new" UEC to the right-hand side, it allowed for an added linear relationship between the new and old UECs. The second additional functional form was also similar to the first but removed the logarithmic transformations. This dependent variable was the difference used in Figure 4 above.

These three basic functional forms were each run four times to account for the difficulties of matching the metered units to manufacturers' data. Minimum and maximum production age (and associated UEC) provided two different data series. Each of these series was run using the more restricted, complete matches and the larger dataset that also included the close matches. In all, then, there were 12 regression models with 12 unique combinations of functional form, data series, and data scope.

Effect of Age Alone

The first step was to confirm the indication from observation of Figure 4 that there was a weak relationship between usage difference and age. We ran all 12 regression models including only the production age (minimum or maximum) and the sample indicator as explanatory variables. Production age was statistically significant at a 90 percent confidence level for only 2 of the 12 regression models, Model #3, using complete matches for both production ages. These regression models showed an increase of 38 kWh/year for minimum production age and 36 kWh/year for maximum production age. The regression fits did not indicate a difference between the two meter samples. These results indicated that using the correct functional form and limiting the data to the complete matches there is in fact a discernible relationship between age and difference in old and new UEC. However, the R^2 of .06 indicated that very little of the variation in the data was being explained.

Effects of Age with Other Characteristics

We explored the possibility that the relationship between usage difference and age was informed by some combination of unit characteristics. Table 1 provides the list of characteristics we tested in the model.

Table 1. List of Variables Explored in the Degradation Regressions

Unit Configuration Binaries	Side by Side
	Single Door
	Top Freezer
	Upright Freezer
Other Characteristic Binaries	Frost Free Binary
	freezer binary
Continuous Characteristics	size
	Amps
	Age (Max or Min Production)
Sample Indicator	Sample Binary (1998 sample)
Amps Interactions	Freezer *Amps
	Below Freezer *Amps
	Side by Side *Amps
Age Interactions	Side by Side *Age
	Single Door *Age
	Top Freezer *Age
	Upright Freezer *Age
	Frost Free Binary *Age
	freezer binary *Age
	size *Age
	amps *Age

We included this list of variables in the models using a backwards elimination model selection procedure. In this procedure, the variable contributing the least to the model is removed until all variables remaining are significant at the 10 percent level. We checked for influential observations and found none.

This approach has its limitations. It can be viewed as "fishing for significance" rather than being based on theoretical justification for a particular model structure. In this case, we are willing to do a little fishing because it simplifies the model saving valuable degrees of freedom. All of the variables we include, for the sake of comparison, have a strong theoretical justification for inclusion in the regression model. With sufficient data, it's likely that all of the variables would remain in the fitted model. As it stands, we have clear indication of an age effect and examples of how characteristics interact with that age effect. Given the data limitations, this is all that can be expected from these models.

Table 2. Variables Remaining and R^2 for All Model Permutations

Variable	Maximum Production Age						Minimum Production Age					
	Close Matches			Complete Matches			Close Matches			Complete Matches		
	#1: Δ log	#2: log v. log	#3: Δ	#1: Δ log	#2: log v. log	#3: Δ	#1: Δ log	#2: log v. log	#3: Δ	#1: Δ log	#2: log v. log	#3: Δ
Side by Side			X									
Single Door												X
Top Freezer										X		
Upright Freezer				X		X						
Frost Free Binary				X	X						X	X
freezer binary				X		X			X	X		
size				X	X	X			X			
Amps												
Age (Max or Min Production)	X	X	X	X	X	X	X	X		X	X	
Sample Binary (1998 sample)												
Freezer*Amps				X					X			X
Below Freezer*Amps	X	X	X	X	X	X				X	X	X
Side by Side*Amps												
Side by Side*Age	X	X		X	X	X				X	X	X
Single Door*Age	X	X	X	X	X	X	X	X		X	X	
Top Freezer*Age	X	X	X	X	X	X				X	X	X
Upright Freezer*Age						X						
Frost Free Binary*Age												
freezer binary*Age	X	X	X		X							
size*Age	X	X	X	X	X	X	X	X		X	X	X
amps*Age												X
R^2	.143	.568	.169	.295	.674	.337	.088	.546	.141	.232	.652	.321

The variables included in the final version of each of the 12 regression model permutations are indicated in Table 2. While there was some variation with regard to which variables remained in the models, aspects of the results were consistent across all the permutations. Age is significantly different from zero in 10 of the 12 permutations, only dropping out for the un-transformed difference models (#3) using minimum production age. Age, interacted with other unit characteristics, enters into all of the models at least four times in most of the models. These results indicate that door configuration, size, and whether a unit is a freezer or not combine with age to explain unit degradation.

The coefficient of age is negative in all 10 instances where it is present. This result seems to suggest that UEC decreases with age. However, in all the models where age was retained, interactions between age and other variables were also retained, and with a positive sign. Thus, to determine the overall effect of age, the net effect in combination with other factors must be considered. The results for the best regression model are discussed below, and Figure 5 shows that the net effect of age remains positive. That is, UEC increases with age.

The results also highlight aspects of the different data we are using for different regression model permutations. Maximum age consistently outperforms minimum age in terms of R^2. Similarly, the data from complete matches consistently performs better than the larger, close matches dataset. The best set of models, the three models using maximum production age and the complete matches, yield similar results. Models #1, 2, and 3 have 11, 9, and 10 characteristic variables included and 7 of them are in common. The two logged models (models #1 and 2), with variables in the same scale, have very similar magnitudes across like variables. The explanatory power of the models appears also to be similar. Model

#2, which includes new usage as an explanatory variable, has a much higher R^2, as would be expected. However, the similar coefficients for the two logged models, models #1 and #2, indicates that model #2 is similar to model #1 in explanatory power with regard to the relationship between "old" and "as new." Model #1's R^2 is just slightly less than that of model #3.

Best Model

Parameter estimates for model #3 using maximum production age and complete matches are reported in Table 3. This model is representative of the three best models with regard to the included age interaction variables. As this model is not log-transformed, the parameter estimates represent annual kilowatt-hours and thus are more intuitive. The parameter estimate for maximum production age for model #3 is negative 455.0, but all the age interaction/binary characteristic terms are positive. Age interacted with size is also positive indicating that larger units degrade at a higher absolute rate (kWh/year) than smaller ones, as would be expected. Combining the effects of age alone with the interaction effects, side-by-side, single door or top freezer refrigerators, and upright freezers above modest sizes (17,14,16, and 6 cubic feet respectively) are all found to have a positive net age effect increasing with size. Upright freezers have the strongest age-related effect.

Table 3. Model #3 Parameter Estimates, Maximum Production Age, Complete Matches

	Parameter Estimate, Annual		Standard Error	Pvalue
Intercept	3,119.4	kWh	1,629.7	0.059
Upright Freezer	-9,253.4	kWh	3,400.4	0.0079
freezer binary	6,944.6	kWh	2,497.1	0.0067
size	-168.9	kWh/Cu. Ft.	82.7	0.0442
Below Freezer	698.4	kWh	213.7	0.0016
Side by Side * Age	232.9	kWh/Year of Age	86.7	0.0087
Single Door * Age	267.3	kWh/Year of Age	88.9	0.0035
Top Freezer * Age	245.1	kWh/Year of Age	87.0	0.0061
Upright Freezer * Age	381.9	kWh/Year of Age	138.2	0.007
Size * Age	13.6	kWh/Year of Age	4.5	0.0033
Maximum Production Age	-455.0	kWh/Year of Age	123.8	0.0004

To assess the implications of these results for the units being collected by the program, we used model #3 to estimate the mean per unit increase in UEC for the mix of refrigerator models in the 2003 tracking. The characteristics reflect the 2003 tracking data, except for age. As tracking age was shown to be a poor estimator of production age, we applied the model multiple times with all units assigned a single age, 5 years, 10 years, etc. Thus, we effectively assume that the distribution of unit characteristics other than age is the same as observed for the population of units collected in the 2003 program, and look at the effect of age on such a mix of units. In reality, of course, characteristics such as size and configuration vary with age due to changing production practices. However, this approach lets us see what the overall age-effects appear to be for the units observed. The same type of analysis could be done for any particular distribution of characteristics of interest.

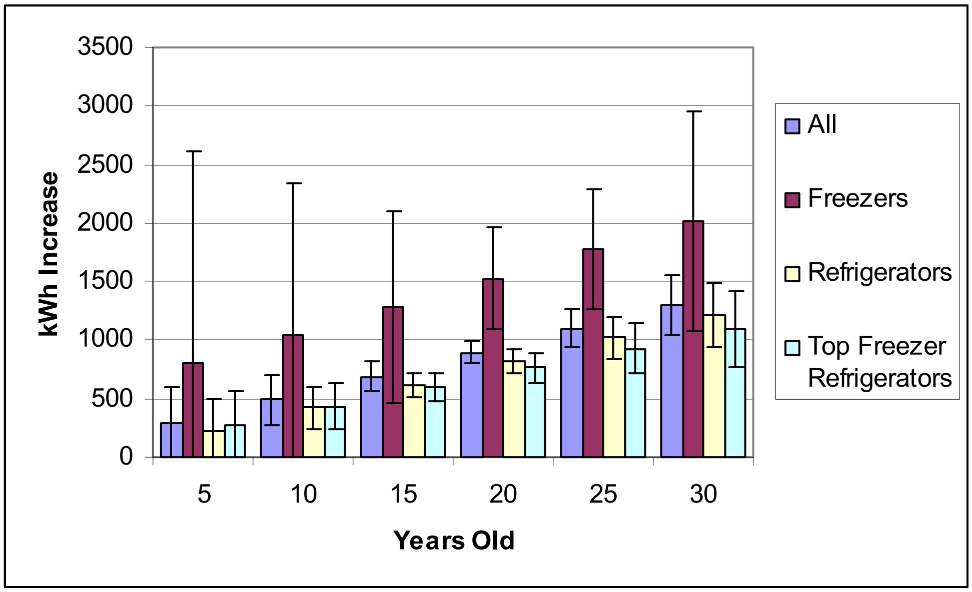

Figure 5. Estimated kWh increase in UEC at Five Year Age Increments

Figure 5 shows the results from model #3 for the full 2003 RARP tracking data and three different subgroups. The figure shows the estimated mean per unit increase in UEC as a function of age, with 80 percent confidence bands, for each of the four subgroups. The overall results show an average increase in UEC of 40 kWh/year. Thus, the regression model including characteristics and applied to the 2003 tracking data population indicates degradation of a similar magnitude to the age-only regression model results reported above The increased R^2 of .34, compared to 0.06 for the age-only regression models, indicates that inclusion of characteristics explains a substantial portion of the variability in the data. Given the lack of data with a production age of less than 10 years, the results for 5-year-old units are not reliable; they are included to illustrate the trend determined by the model fit.

The importance of the inclusion of characteristics in the model is illustrated by the three sets of characteristic-specific results in Figure 5. Freezers clearly show a higher degree of degradation. Because the number of freezers in the analysis is small (7 of the 96 complete matches), the 80 percent confidence band is quite wide. Similarly, the most common refrigerator configuration, those with a top freezer, shows lower degradation than refrigerators as a whole.

Conclusion

This analysis clearly indicates that degradation, in terms of increased UEC, does take place in refrigerators and freezers over their lifespan. Ninety-three percent of the matched metered units showed increased usage during their lifespan compared to the "as new" rating for units of the same model. The median difference was an increase of 557 kWh/year; the median percent difference was an increase of 46 percent.

The regression analysis further indicates that degradation can be quantified as a function of unit characteristics. In particular, the effect of age varies with unit configuration and size.

It would, in principle, be possible to estimate potential savings for a program population using these results. Furthermore, potential savings from any given production year could be determined. With this information, it would be possible to determine a minimum unit age to target to optimize program savings goals.

There are a number of issues with the data that make the specific results reported in this paper tentative. These include:

- Manufacturer's data limited in years, varying over years and interpolated among models,
- Uncertainty regarding the exact age of the metered units, and
- The imprecise nature of the matching of model numbers.

Also, the substantial majority of units used in our models were manufactured prior to 1990. These results are thus primarily illustrative of degradation in pre-1990 units. Whether similar rates of degradation will be found in the post 1990 units, or, put another way, whether units of more recent vintage will age in the same way as earlier ones did, remains a question for future research.

These degradation results are based on a sample of units turned in to the recycling program. As noted, surveys conducted as part of the associated evaluations determined that, without the program, the great majority of these would have been discarded by their owners by some other means. Thus, these degradation results for units collected by this type of program do not necessarily apply to the average unit of the same age currently in place in homes.

Finally, given the small size of our sample and relatively low explanatory power of the best models, these models are of limited use in providing firm quantitative estimates of potential degradation in specific units. A larger sample size, incorporating additional metering cases from studies currently planned, may provide more robust estimates.

Despite these limitations, this analysis provides a strong indication of the kinds of results that may be attainable with an analysis of this sort. With more data and more recent units, results like these could prove invaluable to future recycling program administrators.

References

KEMA-XENERGY Inc. 2004. *Measurement and Evaluation Study of the 2002 Statewide Residential Appliance Recycling Program.* Madison, WI: KEMA-XENERGY Inc.