EXAMINING EXPENDITURES ON ENERGY

Louise A. Heslop[1]

Using data from the Family Expenditures Surveys over time, consumer expenditures on in-home and transportation energy from 1969 to 1982 are being studied. This article briefly summarizes some of the procedures being used to explore the data, summarize it and develop insights into shifts in consumption for policy implications purposes. With such a complex data set and such a complex, multi-faceted subject for analysis some effort must be made to reduce information flows and at the same time increase the information content of each factor of both input and output in the analyses.

1. THE ENERGY ISSUE

To some, energy conservation may be a dead issue. There is no shortage of energy (maybe never was): prices for energy have stabilized.

Energy matters dominated the 1970's having major impacts on the world economic order and creating international strife. Domestically they impacted drastically on federal - provincial relations and business - government relations and on family budgets: caused the restructuring of the manufacturing base, the auto industry, etc. Despite its reported demise as an important issue, energy consumption and prices remain as high priority concerns of consumers, businesses and governments. Energy conservation has lost its sparkle but not its real value.

The research I will be reporting on briefly has been developed in consultation with policy makers in Consumer and Corporate Affairs Canada and Energy, Mines and Resources Canada which continue to run active research programmes on consumer energy use and conservation. The project structure has taken their interests, orientations and limitations into consideration.

Also, within the last five years an international group of social scientists has begun a series of research and information exchanges on consumer behaviour and energy use. As a member of that group I have been keenly aware

[1] Louise A. Heslop, Research and Analysis Division, Statistics Canada.

of the problems and prospects and the current state of knowledge and research techniques of that group.

2. PROBLEMS IN ENERGY RESEARCH

Perhaps the major problem in studies of consumer energy use has been to obtain reasonably reliable measures of use from sufficiently large and representative samples. Getting such data over a period of time, especially a time period spanning the infamous 1973 oil embargo period, would send a researcher into Nirvana. The Family Expenditure data collected by the Consumer Income and Expenditure Division of Statistics Canada come close enough to these requirements to at least set one's heart fluttering. It is a series of retrospective recall studies conducted for the years 1969, 1972, 1974, 1976, 1978 and 1982. So it covers the time period of interest for a large sample and the sampling technique used ensures that the design is representative of Canada for those areas studied, usually urban centres. Additionally it contains a great many other variables of interest in any study of energy use, e.g., home ownership, some house characteristics, vehicle and appliance ownership, family characteristics and expenditures on other categories of consumer goods and services, etc.

Most studies which attempt a measure of consumer expenditures rely on recall or file checking by respondents. There are obvious problems with the accuracy of such data on an individual basis. The problems are less restrictive with very large samples. For most independent studies, the costs of such large samples are prohibitive. However, FAMEX sample sizes are very large.

Only one major study in Canada has used independent record checking, obtaining records from suppliers by household with the permission of the householder, but through this technique was able to obtain electricity use records on less than half of its sample. Natural gas and oil records were obtained on only about one-third of the sample. This procedure of record checking is highly accurate, removes the problems associated with recall, especially over long periods of time, and of reporting bias of respondents. However, practically it is impossible to use for large samples across the country.

Although the FAMEX Study uses recall procedures, the information on energy expenditures are not likely to be as biased as in a study specifically designed to record energy behaviours since respondents are not sensitized to the subject of the study. Also the data from pre-energy crisis periods was collected in the same way as that since the crisis, again reducing the likelihood of response bias. So the FAMEX data set offers a unique opportunity to examine a very large set of samples during a very important period of time.

The data set is not without its problems, some because of the sampling procedure and some because of the inherent complexity of any study of energy use. Changes in expenditure categories and their contents, especially those other than energy, have required that we manipulate the data considerably to create consistency across years. It is not possible to track in-home energy expenditures for those families who do not pay for energy directly, i.e., apartment dwellers with central metering and roomers. Some researchers have imputed values to these households based on their rents but we chose not to, and instead have chosen to restrict our study to those households who have the ability to monitor and affect their own energy use. These households are the consumer groups who will be the focus of any government programmes to alter consumer comsumption.

There are several factors which make the study and the altering of energy consumption of households difficult:

- Capital commitments restrict the ability of the household to respond in the short-term and increase the cost of response - e.g., house size, number and type of appliances, size and number of vehicles. Some studies have noted that home characteristics alone may account for 24% of in-home energy consumption. Family size may be considered as a capital commitment as well.
- Flow feasibilities - There are restrictions in the ability to change the amount and types of fuels used depending on the technology and fuels available under different circumstances and for varying amounts of money, e.g., natural gas heating is not available to rural residents: instantaneous changes can not be made in the type of home heating fuel used.
- Exogenous factors affect the amount of energy needed for similar

performance in different situations, e.g., weather, distances between points in cities, etc.

3. SUMMARIZING INFORMATION INPUTS AND MAXIMIZING INFORMATION OUTPUTS

With such a complex data set and such a complex, multi-faceted subject for analysis some effort must be taken to reduce information flows and at the same time increase the information content of each factor of both input and output. There are several ways of doing this, some of which we will be using, they include:

a) Constructing Complex Input Variables - to reduce the number of factors being studied to the most salient ones.

i) Discontinuous complex input variables were created by combining in-home and transportation energy consumption but not as continuous variables. Rather groupings were created to develop a set of typologies whose characteristics can then be examined for differences. In this case the groupings were developed by creating expenditure quartiles for each energy category, collapsing the two middle categories, and then combining the two resulting three cells into a nine cell matrix of interrelated categories (see Table 1, source: McDougall, Ritchie and Claxton). In particular, the corner cells are of interest in contrast to each other and to the middle cell. This typology was developed in an earlier study for Consumer and Corporate Affairs Canada. So comparing the output from the FAMEX data to the data set used in the CCA study will be of particular interest. Comparing the characteristics of these groups over time will also be of interest. For example, do the Churchmice continue to be impoverished Canadians (involuntary simplicity) or is there any indication that there is some voluntary embracing of low energy, lifestyles? In Table 2 the characteristics of three cells of the typology from two different years are compared - the Churchmice, the Roadrunners and the Hippos. Looking first at the Churchmice, information on a selection of possible analysis variables is shown across two different years, 1974 and 1978. To simplify for this presentation only the rankings of the cell within the typology set of cells is given. Characteristically those consuming the least amount of energy have had the least resources in general, i.e., the lowest incomes, the lowest levels of education, the oldest. These characteristics are evident for the Churchmice in 1974, they also have the lowest levels of consumption for all the expenditure categories shown. Although they are the oldest group they do not have the lowest number of very young children. Probably this group consists of a mix of senior citizens and single parent households (probably headed by women) with young children. Note that this group also has the lowest number of full-time earners (F-T earners). In 1978 the general picture is still the same except that this group is no longer the oldest. In fact the oldest group is in the adjacent cell to the right in the typology (not shown here). It would seem that in 1978 the very old are consuming a relatively larger amount of in-home energy. Perhaps this group is financially better off in 1978 than in 1974 or perhaps they have been unable to hold the line on energy expenditures as prices have risen.

In 1974 the Hippos also fit expectations. They seem to be middle-aged with large numbers of children 5-16 years of age. The "full nest" family, they spend the largest amount on most expenditure categories. They are also the most highly educated. In 1978 this is no longer true as the education ranking of this cell has dropped. Also this group no longer has the highest shelter expenditure. Some suggestions for these observations may be that those with the largest homes and the highest education have begun to modify their homes to reduce energy expenditures.

The Roadrunners have changed also. In 1974 they were the youngest group with very small families. In 1978 they appear to be characterized as young families with young children. One of the most dramatic changes for this group has been that their alcoholic beverages and tobacco expenditures have dropped dramatically.

The significance of these changes can be determined with appropriate statistical tests. The purpose of this discussion was to introduce the idea of searching for meaningful typologies within the data. Pictures of the lifestyles of the groups emerge which can be very useful in furthering conservation programmes directed at each group.

Further analysis may look not at level of expenditures but at percent of expenditures. Such an analysis will reveal the characteristics of those who are most heavily burdened with energy bills.

ii) Continuous complex input variables can be constructed to eliminate the effects of variables known to have very large effects, but ones which are difficult or impossible for consumers to manage.

In-home energy expenditures can be examined for factors related to them, but since one of the main determinants of in-home energy expenditures is house size, this size factor can be absorbed into the input variable to allow for examination of other more relevant (from a policy perspective) factors. So instead of in-home energy expenditures, in-home expenditures/room are examined. Taking this one step further, climate and weather variances from year to year may be controlled for by looking at expenditures/room/degree day. This last factor is added to the data set by city by year. Degree day data for each year for each city were obtained from Environment Canada. Table 3 indicates how the figures change as the factor studied becomes more complex again across two of the years of data. A comparison of the two years and differences in the measures of change between years suggests the importance of refining the measure to improve understanding of the process.

b) Constructing summary output variables to examine the structure of the data - Example of regression coefficients.

In Tables 4-6 some regression outputs are presented. Three models are examined. In each succeeding model the dependent variable becomes more complex. In so doing the factors known to impact significantly on energy consumption can be controlled for and the effects of the remaining variables examined more constructively for any significant explanatory power.

In these analyses no attempt has been made to deal with the problem of the complex sampling design. A future analysis will do so using the Taylor linearization procedure and results will be compared. However, the results from both a weighted and an unweighted sample are shown for 1974. As can be seen the values of the coefficients change very little and their significance or lack thereof does not change. Because of the restrictions indicated and also the fact that the very large sample sizes are used here produce significant results under conditions of very slight differences, it is advised that great care be taken in viewing these preliminary results for purposes of this discussion. I will only note the variables significant at the .01 level and beyond and then only their sign.

In the independent variable list dummy variables are used in the first and second models for city and in all three models for type of dwelling type. The unspecified condition is Ottawa for city and single detached house for dwelling type.

In 1974 house size, some city variables, total expenditures, age of head and family size and some house types are significant. Large families with high total expenditures living in single detached homes in St. John's consume the most. Western cities consume less than the east, and all other housing types consume less than detached houses, although duplexes not significantly so when number of rooms is controlled for. The unweighted results are similar to the weighted.

When the dependent variable is changed to \$/room and number of rooms is removed from the list of independent variables the general pattern remains. However, family size is no longer significant (probably closely tied to dwelling size only), and education of family head becomes significant with a negative sign. Those with less education consumed more, all other things being equal. Finally duplexes become significant with a positive sign, so when number of rooms is controlled for, duplexes use more energy than detached houses.

In model 3 climatic conditions are taken into account by controlling on degree days in the dependent variable and the list of cities is dropped from the independent variable set.

It should be noted that the value of the coefficients drops so dramatically because there are between 4000 and 7000 degree days in these cities. So the small value of the coefficients does not mean they are unimportant. Total expenditures remains significant as does education of the family head and the rowhouse effect. An important thing to note is the drop in the value of the adjusted R-squared. In fact the independent variables remaining in the equation do not do very much to help in explaining variance in the dependent variable. Other more useful variables should be sought.

When we compare just the unweighted 1974 and 1978 results, in model 1 some change in the Vancouver parameter can be noted and in the importance of semi-detached and duplex housing over detached houses.

In model 2 again the major change is in dwelling type effects. Finally in model 3 only the rowhouse variable shows any difference from the detached:

education of the head is again important, but in 1978 age of head is significant with a positive coefficient. Some improvement is seen in the R-squared for 1978, but it is still very low.

This cross-year comparison from a policy perspective suggests perhaps that improvements have been made in the quality of the detached housing stock in Canada. From a methodological perspective it indicates the importance of choosing the dependent variable with care.

As was earlier noted, much additional analysis and re-analysis will be done using the regression procedures available to refine these results and take the sampling design into account.

As I noted earlier the FAMEX data sets have their limitations but they also contain a wealth of important information which should be fruitfully explored.

REFERENCE

[1] McDougall, Gordon H.G., Ritchie, J.R. Brent, and Claxton, John D. (1979). "Energy Conservation and Conservation Patterns in Canadian Households: Overview." Behavioral Energy Research Group, 203-2053 Main Hall, University of British Columbia.

Table 1: Energy Consumption Taxonomy - Labels

		Level of In-Home Energy Consumption			
		Low 127 Mil. kJ	Medium 127-222 Mil. kJ	High 222 Mil. kJ	Total
Level of Automobile Gasoline Consumption	Low 1136 litre	CHURCH MOUSE 4.5% of sample	9.8% of sample	BEAR 2.5% of sample	16.8
	Medium 1136-4545 litre	14.5% of sample	BEAVER 33.7% of sample	12.3% of sample	60.5
	High 4546 litre	ROADRUNNER 4.0% of sample	12.6% of sample	HIPPO 6.1% of sample	22.7
	Total	23.0	56.1	20.9	100.0

Source: See reference list.

Table 2: Rank among Typology Cells

	Churchmice		Hippos		Roadrunners	
	1974	1978	1974	1978	1974	1978
Education of Head (low-hi)	1	1	9	7	7	8
Age (old - yng)	1	2	6	6	9	9
F-T Earners (low-hi)	1	1	8.5	9	7	6.5
Family Size (low-hi)	1	1	9	9	4	4
Child Less than 5 (low-hi)	3	1	4	2	1.5	7
Child 5-15 (low-hi)	1	2.5	7	6.5	5	2.5
Food at Stores (low-hi)	1	1	9	9	4	4
Food at Eating Places (low-hi)	1	1	9	9	6	6
Shelter (low-hi)	1	1	9	7	4	3
Clothing (low-hi)	1	1	9	9	6	5
Personal Care (low-hi)	1	1	9	9	5	4
Medical (low-hi)	1	1	8	8	4	4
Tobacco & Alcohol (low-hi)	1	1	9	9	7	4
Reading, Recreation, Education (low-hi)	1	1	9	8	8	9

Table 3: Average In-Home Energy Expenditures, 1974-78

	1974	1978	% Change
Average $ in-home energy expenditure	451	764	+69
Average $/room in-home energy expenditure	73	121	+66
Average $/room/dd in-home energy expenditure	.019	.029	+53

Table 4: Regression Analysis Results - Model 1 - $In-Home Energy

	1974 Unweighted	1974 Weighted	1978 Unweighted
Intercept	197.3 A	225.4 A	298.0 A
No. of Rooms	13.9 A	12.0 A	4.2 C
City - St. John's	193.9 A	204.9 A	341.1 A
Halifax	75.5 A	73.9 B	162.0 A
Montreal	12.2	22.7	-16.6
Toronto	-10.2	-3.0	50.5
Winnipeg	-127.1 A	-125.4 A	-72.2 C
Edmonton	-244.9 A	-243.2 A	-195.8 A
Vancouver	-22.9	-17.5	-71.9 C
Total Expenditures	.006 A	.006 A	.01 A
Age of Head	1.2 A	0.8 B	3.6 A
Family Size	13.2 A	12.1	21.6 B
Education of Head	0.7	0.6	-3.6
House Type - Semi Det.	-50.9 B	-49.0 A	-23.8
Rowhouse	-81.2 A	-88.9 A	-119.7 B
Duplex	-12.3	-13.7	-84.6 C
Adjusted R^2	0.43	0.34	0.38
F value (prob.)	118.5(.0001)	79.7(.0001)	74.6(.0001)

Note: A = prob. less than .0001, B = prob. less than .001, C = prob. less than .01

Table 5: Regression Analysis Results - Model 2 - $/Room

	1974 Unweighted	1974 Weighted	1978 Unweighted
Intercept	76.2 A	77.3 A	99.8 A
City - St. John's	30.4 A	32.0 A	74.8 A
Halifax	16.8 A	16.3 B	31.6 A
Montreal	4.5	6.7	6.5
Toronto	-3.5	-1.7	10.1
Winnipeg	-17.6 A	-16.3 A	-0.9
Edmonton	-37.9 A	-36.8 A	-26.4 A
Vancouver	0.3	0.8	-6.7
Total Expenditures	2.2×10^{-4} B	2.5×10^{-4} B	6.9×10^{-4} A
Age of Head	0.015	-0.03	0.33 B
Family Size	0.6	0.04	-0.63
Education of Head	-1.9 A	-1.4 B	-4.0 A
House Type - Semi Det.	-6.5 C	-7.1 B	3.1
Rowhouse	-11.5 A	-11.8 A	-11.0
Duplex	6.1 C	6.6 C	3.24
Adjusted R^2	.31	.19	.24
F value (Prob.)	73.85(.0001)	38.9(.0001)	41.4(.0001)

Note: A = prob. less than .0001, B = prob. less than .001, C = prob. less than .01.

Table 6: Regression Analysis Results - Model 3 - $/Room/DD

	1974 Unweighted	1974 Weighted	1978 Unweighted
Intercept	.017 A	.019 A	.02 A
Total Expenditures	8.01×10^{-8} B	9.4×10^{-8} A	1.4×10^{-7} A
Age of Head	1.8×10^{-5}	-7.0×10^{-6}	9.9×10^{-5} A
Family Size	-1.4×10^{-5}	-18.4×10^{-5}	27.0×10^{-5}
Education of Head	-5.3×10^{-4} A	-4.7×10^{-4} A	-7.8×10^{-4} B
House Type - Semi Det.	3.4×10^{-4}	-7.5×10^{-4}	24.8×10^{-4}
Rowhouse	-23×10^{-4} C	-35.9×10^{-4} A	-38.8×10^{-4} B
Duplex	16.9×10^{-4}	6.3×10^{-4}	11.6×10^{-4}
Adjusted R^2	.01	.02	.03
F value (Prob.)	5.6(.0001)	6.6(.0001)	9.5(.0001)

Note: A = prob. less than .0001, B = prob. less than .001, C = prob. less than .01