CBEEDAC
CANADIAN
Building Energy End-Use
DATA AND ANALYSIS CENTRE
commercial • residential • institutional

Explaining Energy Savings under the EnerGuide for Houses Home Retrofit Program

David L Ryan*

May 2009

CBEEDAC 2009–RP-02-DRAFT

---

[*] The financial support of Natural Resources Canada through the Canadian Building Energy End-Use Data and Analysis Centre (CBEEDAC) is gratefully acknowledged. Earlier versions of this paper were presented at the 31st International Conference of the International Association for Energy Economics, Istanbul, June 2008, and at the 1st International Workshop on *Empirical Methods in Energy Economics*, Zurich, Switzerland, August 2008. I am grateful to Denise Young, Stuart Landon, and Mark Maxson for very helpful discussions, and also to Mark Maxson, and earlier Sam Gamtessa, for research assistance.

Executive Summary

Data collected in two rounds of residential energy audits under the Canadian EnerGuide for Houses (EGH) Program are used to model and empirically investigate the determinants of the decision to retrofit for those that completed the first audit, as well as the factors that affect the energy savings that were achieved by homeowners who completed the second audit. The model formulation accounts for the fact that the EGH energy consumption data used in the estimations are inferred from technological relationships rather than reflecting actual measures. Thus, estimation of a model specification where energy savings depends on changes in the energy-efficiency characteristics of the house along with behavioural (household) factors would simply result in an imprecise estimate of the technical relationship that was used to infer the energy consumption values, with the behavioural factors appearing to play no role. To avoid this problem, as well as endogeneity of the energy efficiency characteristics of the house following the retrofit, an alternative model specification is developed. Since the EGH data do not contain household information, census variables are matched to the EGH data using part of the postal code. Alternative estimation approaches using the model that is developed, which take account of the endogeneity of the decision to retrofit, reveal the importance of behavioural factors in both the decision to retrofit and the energy savings that are achieved. Knowledge of the relative importance of these types of factors could help better target subsequent programs of this type so that greater energy savings are achieved.

Table of Contents

Executive Summary …… i
List of Tables …… iii
List of Figures …… iii

1. Introduction …… 1
2. The Canadian EGH Program …… 4
3. Models and Estimation Methods …… 12
4. Results and Analysis …… 21
    4.1 Probability of Retrofitting (and Completing the Second Audit) …… 28
    4.2 Explaining Energy Savings …… 32
5. Summary and Conclusions …… 36

References…… 39

List of Tables

Table 1: Number of Homes in the EGH Data Set: First and Second Audits ......…………... 6
Table 2: Upgrade Categories by Vintage ………………………..……………………….... 7
Table 3: Upgrade Intensity and its Determinants ……………………..………..................... 9
Table 4: Average Post-Retrofit Energy Savings and pre-Retrofit House Characteristics ….. 11
Table 5: Energy Intensities ……………………………………………………………. 23
Table 6: Details of Variables Included in the Analysis …………………………………….. 24
Table 7: Marginal Effects from Probit Estimation of (17) …………………………………. 30
Table 8: Marginal Effects from Tobit, Truncation, and Sample Selection Estimations ……. 33

List of Figures

Figure 1: Number of Homes with Specific Upgrades ……….........................…………... 8
Figure 2: Pre-Upgrade Energy Intensities …………………...……………………………... 10
Figure 3: Distribution of Energy Savings …………………………. ……….................... 23

.

1. Introduction

In the residential sector in Canada, energy is primarily used (almost 60%) for space heating purposes, with the next major uses being water heating and lighting. In view of the significant share of this sector in total energy consumption, and hence in greenhouse gas emissions, policy analysts have long touted the need for increased residential energy efficiency. Of course there are several ways that this increased energy efficiency can be achieved. These include imposing higher energy efficiency standards on new buildings and appliances; programs to increase consumer awareness of their energy consumption, GHG emissions, and measures that they can take to reduce them; and programs to encourage residential retrofitting to improve the thermal qualities of a residence as well as the energy efficiency of major appliances such as space and water heaters that are in place. The focus of this report is on residential retrofitting programs and their effect on energy consumption.

While these measures all appear to be obvious ways to increase energy efficiency in the residential sector, recent research has shown that technological (engineering) estimates of the energy that will be saved through introducing products that embody newer energy-efficient technology are not always fully realized. In part this may be due to rebound effects, where increased efficiency is offset to some extent by induced higher levels of usage. However, it can also be due to many other factors, such as less than full utilization of the energy-saving features of the new technology (such as with programmable thermostats that are not programmed). In other cases over-estimates of energy saving associated with new technology occur because the take-up of the new technology is slower than expected, possibly due to costs associated with the new technology that were not anticipated but which become widely known, such as reliability issues, higher costs of replacement parts, and even aesthetic concerns (such as with compact fluorescent lights). Whatever the reasons, the net effect is that the energy that is saved is less than anticipated, and in order to understand the extent to which this is likely to occur, it is necessary to take account of behavioural factors, that is, the fact that the actual energy savings will be the result of the interaction between human behaviour and the change in the energy efficiency properties of the new technology.

A difficulty with empirically determining the actual energy savings by taking account of human behaviour is that the required sample information is often not collected. In the particular example that motivates the analysis in this paper, the Canadian EnerGuide for Houses (EGH) program, all the data that are obtained are based on technological information. Under this program, homeowners voluntarily participated at their own expense (although in some cases with a subsidy) in an initial home energy audit and were subsequently provided with a set of energy-saving retrofit recommendations. They were free to undertake as few or as many of these as they wished, and subsequently could voluntarily undertake, again at their own (in this case considerably lower) expense, a second home energy audit. Based on the home energy auditor's assessment of the energy saved as a result of their home retrofitting, the homeowner was then eligible for a grant that was an increasing function of the energy that was saved. The data that were collected during these two home energy audits forms the EGH database. While these data are very detailed on house characteristics, including their general locations, almost no information concerning the house occupants was collected. In fact, even the energy consumption data, and the change in energy consumption measure that is used to determine grant eligibility and the size of the grant, are based purely on technological considerations. No attempt was made to collect actual energy consumption data from the house occupants, so that behavioural factors appear to play no role in any explanation of achieved energy savings.

In view of this feature of the dataset, it is not possible to use the EGH data to directly calculate the *actual* energy savings associated with home retrofitting activities under the EGH program, nor to determine the effect of behavioural factors on these savings. Nevertheless, as we demonstrate in this paper, the available EGH data, along with supplementary demographic and socio-economic information which we obtain from the Canadian census and match by location to households that undertook retrofits as part of the EGH program, can be used to model and subsequently estimate the roles of various socio-economic and demographic factors, as well as house characteristics, in household energy efficiency decisions. To do this we begin by noting that the energy consumption measures for each house contained in the EGH dataset (one from each audit) can only depend on house attributes, and in particular cannot depend on characteristics and behaviour of household occupants (perhaps other than a few simple measures like household size), as this information is not used by the energy auditors in determining the *estimated* amount of energy consumption. Indeed, this is true for energy consumption based on the initial audit – which depends on pre-retrofit house characteristics, energy consumption based on the follow-up audit – which depends on post-retrofit house characteristics, and therefore the difference between these measures, which represents energy savings, and which depends on the difference in pre- and post-retrofit characteristics. However, the homeowner *decisions* of whether to participate in the audit, whether to make any retrofits, and if so how much retrofitting to undertake, will all depend on behavioural factors.

Although there is no information on the actual energy saved via the retrofits, this information would not have been available to the homeowner at the time they made the retrofits anyway. Rather, their primary source of information on expected energy savings would likely have been the technological energy-savings information associated with various retrofit recommendations that were contained in their (first) energy audit, or provided with particular retrofit items that they installed (such as insulation, furnace, water heater, etc.). Consequently, as we demonstrate in this paper, the difference between estimated energy consumption at the first and at the second audit – although itself a purely technological measure – can alternatively be specified as a function of all the factors that resulted in the homeowner making the selected retrofits, that is, household characteristics as well as the efficiency characteristics of the house *prior* to the retrofits. Further, provided the changes in the energy-efficiency characteristics of particular retrofit items that were installed are not included as explanatory variables, estimation of such a function describing the energy savings achieved through retrofitting will no longer be tantamount to simply recovering an imperfect estimate of the formula used by the energy auditors to assess energy consumption. Rather, even though technological relationships (formulas) were used to determine energy consumption at each audit, and hence energy savings, the estimated model will yield potentially useful information on the roles of different factors in determining household energy savings.

Of course, the decision to undertake any retrofits, like the decision of how much retrofitting to do, is endogenous, and this needs to be taken into account in the estimation procedure. We do this by using results from the literature on treatment effects and sample selection. Specifically, we estimate an equation describing the probability of undertaking retrofits, and use estimates based on this equation as instruments in our subsequent estimation of the roles of both house characteristics and occupant characteristics in determining energy savings through retrofits. This type of information, concerning the interaction of household and house characteristics in determining such energy saving, would be expected to be particularly useful for policy makers and energy analysts in their efforts to better target, and assess the extent of the energy savings that are likely to actually be achieved through the use of, various programs like EGH that are designed to encourage increased residential energy efficiency.

The plan of the remainder of this paper is as follows. In the next section we discuss the EGH data and the energy savings that were calculated (inferred) for those who undertook retrofits and completed the second home energy audit. The models that we estimate using the EGH data are developed in Section 3, which also contains a discussion of the estimation methods that are used and their application to our models and dataset. Estimation results are presented in Section 4, while Section 5 concludes.

2. The Canadian EGH Program

Energy-saving home retrofitting has long been touted as "low-hanging fruit" in the drive to increase energy efficiency and reduce both energy consumption and greenhouse gas emissions. Technological improvements that have occurred in recent years have meant that houses built and equipped in the last decade are generally much more energy efficient than those that were built earlier. While it is unrealistic to expect rapid replacement of this older energy-inefficient housing, there are many changes that can be made in existing residential construction for which the associated capital costs appear to have relatively short payback periods in terms of the energy savings that they will elicit.

Despite this apparent cost-effectiveness of many home retrofitting activities, homeowners appear generally to have been relatively reticent to embrace the concept of home retrofitting in a major way. As a result, governments in many jurisdictions have introduced measures that are designed to encourage such activities. In Canada, the federal government introduced the EnerGuide for

Houses (EGH) program, which provides a home energy-use evaluation service to homeowners along with recommendations on energy-efficiency improvements that could be made to their homes. This home energy audit typically costs the homeowner between \$300 and \$350, although the effective cost varies across regions because some provinces subsidize this service. Although it is believed that this cost as well as the costs of undertaking the recommended retrofit activities are recoverable through energy savings, as an incentive for undertaking the recommended retrofit actions, beginning in Fall 2003, grants were provided to homeowners who completed energy efficiency retrofits based on EGH advisor recommendations.[1] The grant amount depended on the difference between the pre- and post-retrofit EGH ratings of the houses, and thus required homeowners to request an initial home-energy audit and undertake sufficient energy-savings retrofits that would be revealed in a second follow-up audit that had to be completed within a specified time limit.

The EGH audit reports have resulted in a very rich data set which contains information compiled during both the first and the second audits (Blais et al., 2005). As of September 2005, 188,368 houses from across Canada had completed the first of the two audits that are required under the EGH program. As shown in Table 1, approximately 32% of these are in Ontario; 19% are in Alberta, and 17% are in British Columbia, regions which have quite different weather and hence energy-use patterns. The fourth largest location for the initial audits is in Quebec (12%), with all other provinces and regions totalling 20%. An earlier analysis of the data available at the end of the year 2000 by Aydinalp et al (2001) showed that the number of houses audited at that time was only 20,000, although this number had grown to 122,723 by July 2004 (Zyniewski and Mistry, 2005).

Fewer than 20% of homeowners who undertook the first audit also completed the second evaluation, although this still totals over 35,000 houses. This number has grown substantially from the corresponding figure observed a year earlier – as of July 2004 only 8,869 homeowners had completed both evaluations, which is less than 8% of the houses that had undergone the first evaluation by that time. This growth in the percentage of homeowners who have had their

---

[1] The EGH program was discontinued in 2006 following a change in the federal government, but was subsequently reintroduced in a slightly different form and with a new name. All data and discussions in this paper refer to the original program.

homes evaluated for the second time – that is, those who undertook sufficient number of recommended retrofits – may be partially due to the introduction of the incentive package for the EGH program in October 2003.

Table 1: Number of Homes in the EGH Data Set: First and Second Audits

House Region	Number of Homes in the first audit	Percent of total	Number of Homes in the Second Audit	Homes in the Second Audit as a percent of the number in the first audit	Homes in the Second Audit as a percentage of the Total Number of Homes in the second audit
Alberta	34858	18.5	7929	22.7	22.5
British Columbia	32573	17.3	6456	19.8	18.3
Manitoba	12668	6.7	2782	22.0	7.9
New Brunswick	1783	0.9	334	18.7	0.9
Newfoundland	2088	1.1	187	9.0	0.5
Northwest Territories	364	0.2	29	8.0	0.1
Nova Scotia	5327	2.8	696	13.1	2.0
Nunavut	64	0.0	0	0.0	0.0
Ontario	60424	32.1	11343	18.8	32.1
Prince Edward Island	533	0.3	41	7.7	0.1
Quebec	22134	11.8	1581	7.1	4.5
Saskatchewan	13823	7.3	3759	27.2	10.7
Yukon Territory	1729	0.9	152	8.8	0.4
TOTAL	**188368**	**100.0**	**35289**	**18.7**	**100**

In the first audit, information was collected on various aspects of the house and its energy-using characteristics, including details of the major energy-using equipment in the home (space heating, water heating, etc.) and other energy-efficiency features (insulation, windows, etc.). Based on this information, and some basic assumptions regarding household size and thermostat settings (weather), energy consumption data were inferred using technologically-based calculations (formulas) rather than being constructed from actual household energy usage records. Also, almost no information was collected on home owners or occupants. Most of the houses that undertook the first audit were provided with upgrade recommendations by the energy auditors. While homeowners could undertake as few or as many of the recommendations as they wished, the size of any grant that they might receive following the second audit depended on the amount of energy saving that this audit revealed. As part of the second audit, similar information to the first audit was obtained, along with details of the specific upgrades that were undertaken, and post-retrofit energy consumption and energy efficiency was again estimated. These estimates can be compared to the pre-retrofit case as well as to what would have been expected based on adoption of each of the recommended upgrades.

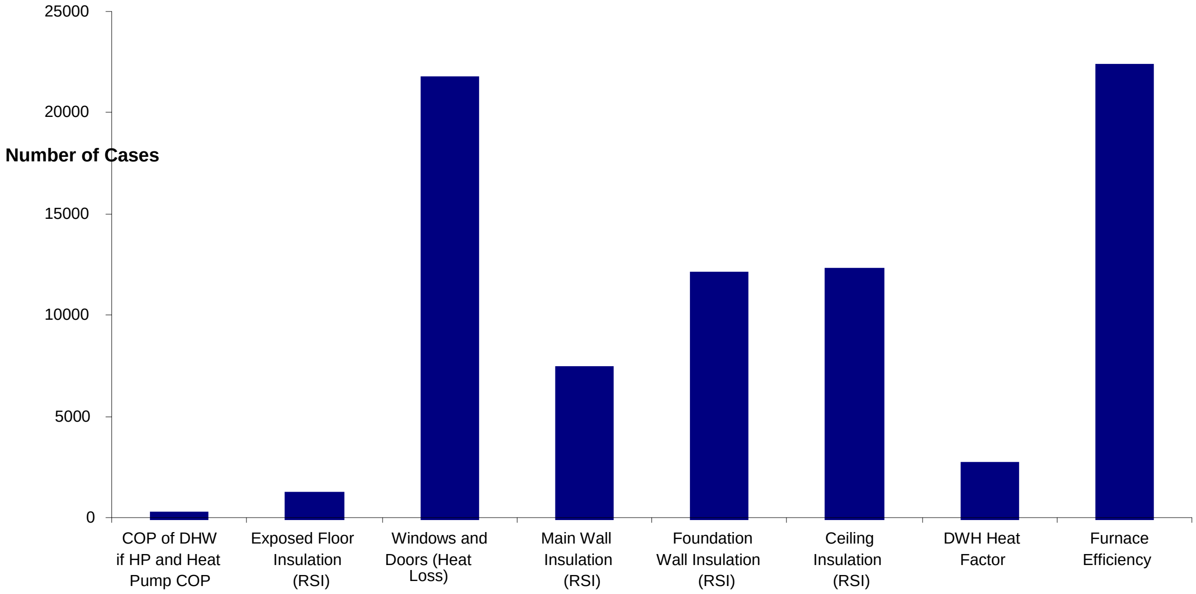
As shown in Table 2, the most widely chosen upgrade type was an improvement in the thermal envelope. For each house age group, the percentage of houses that upgraded their space heating and/or domestic water heating (DWH) system is always considerably less that the percentage that upgraded the thermal envelope. However, as we move from older to newer house age categories, apart from the most recent, the percentage of houses choosing to upgrade the thermal envelope dwindles while the percentage that upgrade their heating system increases.

Table 2: Upgrade Categories by Vintage

Year Built	Heating System (%)	DWH system (%)	Thermal Envelope (%)
1945 or before	19.9	4.1	75.9
1946-60	26.3	3.2	70.5
1961-70	30.6	3.2	66.3
1971-80	31.7	3.0	65.3
1981-90	39.0	4.5	56.5
1991-2005	23.1	6.8	70.1

Of all the possible upgrades that could be made to the thermal structure, as shown in Figure 1, window and door upgrades were the most common, followed by foundation wall insulation and ceiling insulation. This figure, along with the preceding tables indicates that there are considerable differences in the retrofit options chosen by different homeowners, and while in part this may have reflected the energy auditors' recommendations, it also likely reflects decisions made by homeowners based on budget constraints, what they considered to be more important, and the state of their house at the time of the first energy audit.

Figure 1: Number of Homes with Specific Upgrades



As shown in Table 3 below, we observe that houses with greater energy saving potential are more likely to undertake retrofit upgrades, and that the higher is the energy-saving potential, the greater is the number of upgrades that are chosen. In particular, a poor EGH rating generally corresponds with larger number of retrofit upgrades being undertaken, while larger ratios of pre-retrofit to post-retrofit energy consumption and cost (indicators of energy consumption and cost saving potential) generally reflect homeowners undertaking more upgrades.

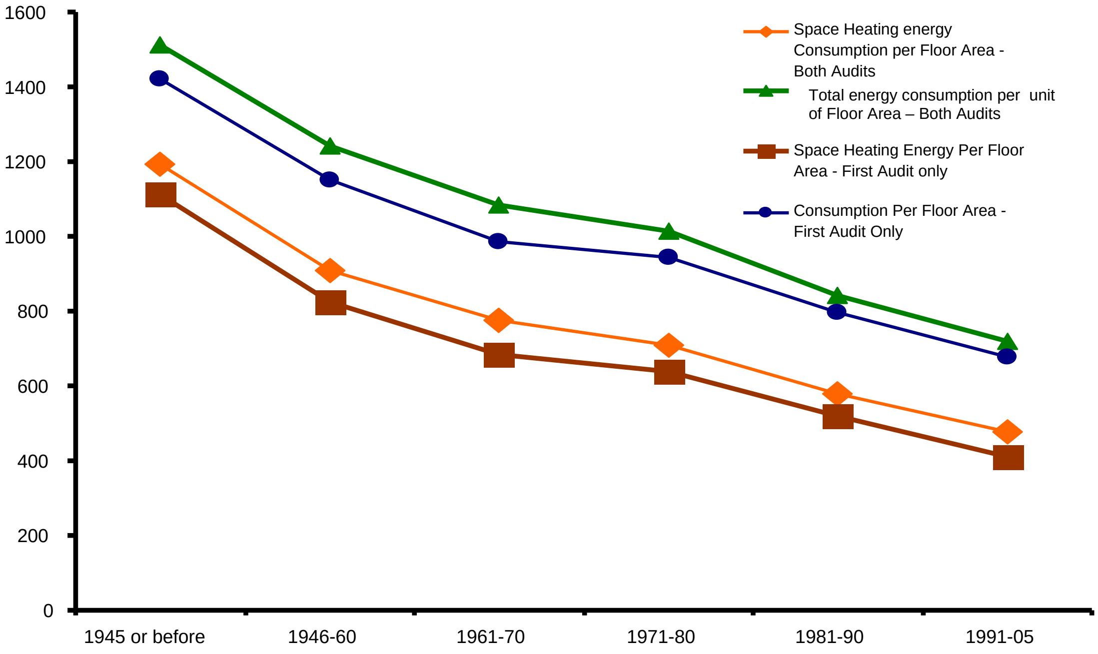
Table 3: Upgrade Intensity and its Determinants

Total Number of Upgrades	EGH Rating at 1st Audit	Energy Consumption Ratio (A/B)	Energy Cost Ratio (A/B)	Blower Door ACH @ 50Pa Ratio (A/B)
0	61.9	1.15	1.16	1.22
1	59.7	1.26	1.19	1.15
2	57.3	1.32	1.22	1.21
3	54.4	1.44	1.30	1.28
4	50.4	1.60	1.42	1.40
5	45.2	1.87	1.57	1.55
6	39.7	2.19	1.70	1.60
7	32.7	2.49	2.00	2.02

Note: "A" refers to the value at the first audit, and "B" refers to the value at the second audit.

This feature – that houses with greater energy saving potential are more likely to undertake retrofit upgrades – is also apparent when we compare houses that did undergo both audits with those that did not. As Figure 2 shows, houses that completed both audits (undertook some energy-saving retrofits) are generally characterized by higher pre-upgrade average energy intensities. This also suggests that the larger the energy and cost saving potential, the higher is the probability that the homeowner invests in retrofitting. Similarly, it also suggests that energy saving is one of the main drivers of retrofit upgrades.

Figure 2: Pre-Upgrade Energy Intensities



Given this pattern, it is interesting to calculate how much energy is estimated to have been saved (based on a technological relationship) as a result of the retrofit upgrades that were undertaken. As shown in Table 4 below, the average energy saving per retrofitted house per year ranges from 35% of pre-retrofit energy consumption (in New Brunswick) to 22% (in Prince Edward Island and Quebec). Of all houses that undertook energy-saving retrofits and completed the second audit, the greatest energy saving per house per year amounted to 88.3% of pre-retrofit consumption. For all provinces combined (last row of Table 4), the average energy savings is about 64,241 MJ per retrofitted house per year, which represents approximately 26% of pre-retrofit consumption. Of course, it is not possible to generalize these observed savings to other houses that did not undertake the second audit. We would expect that houses with better efficiency characteristics would tend to have less energy-savings potential, so that even if they were to undertake various energy-saving upgrades, the energy actually saved would be relatively small. In fact, in many cases it is likely that houses with better efficiency characteristics would not consider it worthwhile to undertake any energy-saving retrofit upgrades.

Table 4: Average Post-Retrofit Energy Savings and Pre-Retrofit House Characteristics

Province	Retrofit Energy Savings (MJ/House per year)	Energy Saving as a proportion of post-retrofit consumption	Average Year Built	Furnace Steady State Efficiency	Primary Domestic Hot Water Energy Factor	Average Ceiling Insulation (RSI)	Average Main Wall Insulation (RSI)	Number of Occupants	Main Floor Thermostat Set-Point
Alberta	64868.48	0.2497	1967.78	72.1123	0.5448	4.3653	1.9099	3.0233	20.985
British Columbia	55859.16	0.2672	1967.43	77.4902	0.5675	3.8208	1.7339	3.6855	20.977
Manitoba	67097.46	0.2555	1954.36	78.5347	0.6082	4.1297	1.7378	3.1334	20.976
New Brunswick	94271.08	0.3471	1951.46	87.4260	0.8071	3.4178	1.7329	3.4671	20.920
Newfoundland	67164.46	0.2580	1958.81	89.7487	0.7540	3.0993	1.8194	3.0695	20.850
Nova Scotia	96839.55	0.2979	1938.17	79.7247	0.6213	2.3081	1.3726	3.7241	20.980
Ontario	62874.61	0.2498	1950.31	79.0504	0.6060	3.6100	1.5610	3.4007	20.951
Prince Edward Isl.	66412.95	0.2178	1948.76	77.8073	0.4949	2.7288	1.9134	2.5122	20.717
Quebec	47911.93	0.2183	1954.11	88.7684	0.7789	3.6218	1.9064	2.9171	20.622
Saskatchewan	76984.56	0.2889	1963.50	70.2690	0.5588	4.6855	1.9864	3.7210	21.013
CANADA	64241.21	0.2580							

3. Models and Estimation Methods

As outlined earlier, the purpose of this paper is to use the available EGH data, along with supplementary demographic and socio-economic information, to model and subsequently estimate the roles of various socio-economic and demographic factors, as well as house characteristics, in household energy efficiency retrofit decisions. As noted earlier, an important characteristic of the EGH data that must be accounted for in any model specification is that energy consumption information contained in this dataset is not observed data, but rather is estimated using a technological relationship. Since this relationship only uses information about the energy efficiency and energy-using features of the house – along with assumed values of weather influences, household size, etc. – to determine energy consumption, this measure obviously cannot depend on the characteristics and behaviour of the occupants of the house.

In view of this aspect of the EGH dataset, we can write energy consumption of the ith household (house) prior to undertaking any retrofits under the EGH (or any similar energy conservation) program, denoted by EC_{0i}, as:

$$(1) \qquad EC_{0i} = f(x_{0i}, w_{0i}) .$$

The explanatory variables in (1) include x_{0i}, which is a vector of variables that represent the energy-using properties of the ith house prior to the EGH program – thermal efficiency, etc. – and w_{0i} which is a vector of other factors that potentially affect energy consumption – household size, weather, etc. Usually, we would think of income and other socio-economic variables as having important roles in determining observed household energy consumption, but since EC_{0i} is just based on a technological relationship, these other variables do not appear in (1).

In the context of the EGH program, the data that are available to estimate a stochastic form of (1) would include households that undertook the first of the two energy audits. As noted earlier, only some of these households undertook energy-saving retrofits and proceeded to the second audit. Of course, the particular houses that undertook the first audit are not necessarily a random sample from the population of houses in Canada. Typically, we might expect that houses that have poorer energy-efficiency properties, or homeowners that are experiencing high energy consumption and are considering retrofits, or even homeowners that are particularly energy conscious – whether or not they have already introduced many energy-savings retrofits and/or appliances in their homes – would be more likely to undertake the first EGH audit. This non-random sample selection could introduce sample selectivity bias in subsequent estimations that use this EGH dataset. Unfortunately, we do not have similar energy consumption and house characteristic information available for non-participants, so that it is not possible to incorporate a sample selectivity correction for this initial non-random selection in our subsequent empirical analysis.[2]

Although participation in the first EGH audit might be considered as participation in the EGH program, the major focus of our analysis is on houses in which retrofits were actually undertaken and a second audit completed. Thus, for our purposes, the "treatment" that we are investigating is completion of the second EGH audit. Thus we define a dummy variable *D*, where

$$D_i = \begin{cases} 1 & \text{if the ith house participated in the } 2^{\text{nd}} \text{ EGH audit} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Since households that have undertaken the 1st audit choose whether to participate in the 2nd EGH audit, presumably on the basis of having undertaken sufficient energy-savings retrofits – which in turn depends on their decisions of whether to undertake any retrofits, and if so, how much retrofitting to do – this variable is endogenous.

Next, for the houses in our sample (all of which undertook the first EGH audit), we define energy consumption by the i$^{\text{th}}$ house after undertaking any retrofits as EC_{1i}. As noted earlier, houses in the EGH program only had a fixed time in which to make any retrofits under the program. Thus, EC_{1i} refers to energy consumption at the end of this period. For houses that did not take any retrofits, which we view as those that did not undertake the second audit $(D_i = 0)$, $EC_{1i} = EC_{0i}$. Like EC_{0i}, EC_{1i} is based on a technological relationship that accounts for the major energy-using appliances in the house, the thermal properties of the house, and assumed values of weather influences, household size, etc. Thus,

---

[2] Alternative ways of dealing with this issue by utilizing information from related datasets, or endogenous stratification techniques are currently being investigated.

(3) $$EC_{1i} = f(x_{1i}, w_{1i})$$

We treat *w*, the vector of other factors that potentially affect energy consumption – household size, weather, etc. – as unchanged between the two audits, so that $w_{1i} = w_{0i}$ for all houses. For homeowners who undertook the second audit, and hence made retrofits $(D_i = 1)$, at least some of the values of the vector of variables that represent the energy-using properties of the house after the retrofits, x_{1i}, will differ from their values prior to the retrofits, x_{0i}. However, for households that did not undertake any retrofits, $x_{1i} = x_{0i}$.

At this point, in order to motivate the model formulation used subsequently, it is useful to consider the type of model that might have been used if we had available *actual* energy consumption data for the houses in the sample. In a typical treatment effects context, data would only be available from a single home energy audit, where some houses in the sample would have participated in the EGH program (and made retrofits, so $D_i = 1$) and some would not $(D_i = 0)$. Generally, the objective of a standard treatment effects model that might be used with this type of data is to determine the effect of program participation on the variable of interest, which in this case is energy consumption. Hence, using the notation developed so far, the model would be written as:

(4) $$EC_{1i} = f(x_{1i}, w_{1i}, h_{1i}, D_i),$$

where h_{1i} is a vector of household characteristics – demographic and socio-economic variables, etc. – which would appear here since these types of factors potentially affect observed energy consumption, in contrast to the case (with the EGH data) where EC_{1i} is calculated using a technological formula. Of course, the variable D_i is endogenous, and estimation methods that are used with treatment models take this into account. However, in the energy consumption context there would be an additional endogeneity problem that would arise with this type of model. Specifically, for households that made retrofits $(D_i = 1)$, the values of the elements of the vector of variables, x_{1i}, that represent the energy-using properties of the house are endogenous, since the decision to participate in the program and retrofit the house necessarily induces changes in the elements of this vector.[3] Consequently, estimation of this model would be complicated by the need to account for this endogeneity.

Even with data from two audits for households that participated in the EGH program, and from only the first audit for those that made no retrofits, estimation of the treatment effects model in (4) would still be feasible, simply by noting that for houses in which no retrofits were made, $EC_{1i} = EC_{0i}$ and $x_{1i} = x_{0i}$. Of course, if actual energy consumption data were available from both audits for houses that made retrofits, there would be little point estimating (4) to determine the effect of program participation on energy consumption, since this could be calculated directly from the data. Even when the energy consumption data that are available are inferred rather than actual, such as with the EGH program data, there is still no point in estimating (4), since in view of the technological relationship that is used to determine energy consumption, as defined in (1) and (3), neither h_{1i} nor D_i would have a role to play in (4) – they would have coefficients of zero. Thus, in this case, estimation of (4) would be tantamount to trying to recover the technological relationship that was used in the first place to determine the energy consumption values that were recorded in the dataset.

In view of these issues, an alternative approach to the development of a model that can be used with the EGH data to estimate the roles of various socio-economic and demographic factors, as well as house characteristics, in household energy efficiency retrofit decisions is required. To begin, we focus on the vector of variables that represent the energy-using properties of the house after the retrofits, x_{1i}. For homeowners who undertook the second audit, and hence made retrofits $(D_i = 1)$, at least some of the values of the will differ from their values prior to the retrofits, x_{0i}. Thus, we can write:

$$(5) \qquad x_{1i} = x_{0i} + \Delta x_i$$

so that substitution into (3), and noting that $w_{1i} = w_{0i}$, yields:

---

[3] In a typical treatment effects context, such as an examination of the effects of participation in a job training program on earnings, for example, the act of participating typically does not affect the values of the other explanatory variables included in the earnings model, such as age, education, etc.

(6) $$EC_{1i} = f(x_{1i}, w_{1i}) = f(x_{0i} + \Delta x_i, w_{0i}),$$

where Δx_i is the vector of changes in the energy efficiency characteristics of the i^th^ house (e.g., furnace efficiency, amount of insulation, water heater efficiency, etc.). For houses that did not undertake the 2nd audit ($D_i = 0$), it is assumed that no retrofitting was done, so that $\Delta x_i = 0$, and hence $EC_{1i} = EC_{0i}$.[4]

Next we consider Δx_i. This represents the energy-saving changes that homeowners make to their house after the first audit. While a number of upgrade recommendations are provided at the completion of the first audit, a homeowner is free to make as few or as many of these as they wish. Consequently, as noted earlier, Δx_i is endogenous. Of course, the actual changes that are made are constrained by budgets and other factors, and in general we do not observe downgrading of these efficiency measures. To model this, we define Δx_i^* as the desired amount of energy efficiency upgrades that the ith house wishes to make, and specify that we only observe this desired amount of retrofits if it is positive. Thus:

(7) $$\Delta x_i = \begin{cases} \Delta x_i^* \text{ if } \Delta x_i^* > 0 \\ 0 \text{ otherwise} \end{cases}$$

When deciding what upgrades to make (if any), a homeowner will (potentially) base their decision on the existing energy consumption of their house as calculated as part of the first audit (EC_{0i}) – which depends on the existing state (thermal efficiency, etc.) of their house (x_{0i}) and other factors affecting energy consumption (w_i) – as well as their financial situation, preferences, etc. These latter factors are captured by a vector of household characteristics – demographic and socio-economic variables, etc. – denoted here by h_{0i}. In addition, the changes they make to their house (and especially the extent of any such changes) might be expected to depend on their decision to participate in the EGH program by having a 2nd audit and (potentially) becoming eligible for a partial grant payment. For example, a homeowner might be

---

[4] Of course it is possible that some of these other houses did some retrofitting but not enough to justify the cost of the 2nd audit. We plan to incorporate this possibility in future analysis.

deciding to replace their old furnace, but because of the program, choose to install a high-efficiency rather than a medium-efficiency unit, thereby increasing their potential EGH grant as well as the energy savings they should achieve from the retrofit. Thus, the actual change in efficiency that they achieve through retrofitting may depend on the existence of and their decision to participate in the EGH program, that is on D_i. Thus, using (1):

$$(8) \qquad \Delta x_i^* = g(EC_{0i}, h_{0i}, D_i) = g(f(x_{0i}, w_{0i}), h_{0i}, D_i) = \tilde{g}(x_{0i}, w_{0i}, h_{0i}, D_i)$$

The model comprising (7) and (8) is a Tobit model that could be estimated directly given data on Δx_i (which are available for those who did the 2nd audit and are assumed to be zero otherwise) and the explanatory variables. However, as noted earlier, Δx_i is a *vector* of changes in the energy efficiency properties of the house, and D_i is endogenous, so that estimation of this model would require simultaneous estimation of several Tobit equations (depending on the dimensions of x_{0i}), each with an endogenous right-hand side binary variable. Also, while it is the case that (by definition) at least one element of Δx_i is non-zero when $D_i = 1$, some elements of Δx_i may be zero even though $D_i = 1$. Thus, for any particular dimension of the retrofit, $\Delta x_i = 0$ does not correspond to $D_i = 0$. However, since $D_i = 1$ whenever at least one element of Δx_i is positive, then the coefficient on D_i will not be identifiable separately from an intercept. Since the dependent variable is the *change* in the energy efficiency properties of the house, it might be expected that an intercept would not normally appear in the function $\tilde{g}(.)$, so that the coefficient that is obtained when an intercept term (or D_i) is included could be interpreted as the effect of the EGH program on the change in (that element of) the energy efficiency properties of the house.

While estimation of (7) and (8) would be of considerable interest, with the estimated parameters revealing the effect of existing house characteristics as well as of homeowner characteristics on the amount of retrofitting that is done with respect to each of the energy efficiency properties of the house, it would be difficult to estimate the simultaneous system, and potentially difficult to aggregate the overall findings in any meaningful way. Fortunately, as described below, both these problems can be remedied by making use of information on the difference between total energy consumption after the retrofits, EC_{1i}, and total energy consumption at the time of the first audit, EC_{0i}. Even though these consumption values are calculated at each audit using a technological relationship, the difference between them essentially aggregates the effects of the changes that were made to all the different energy efficiency characteristics of the house as a result of the retrofitting. Thus, to make the model operational it is just necessary to respecify the model in (7) and (8) to make use of this information.

Corresponding to Δx_i^*, the desired amount of energy efficiency upgrades that the i$^{\text{th}}$ house wishes to make, there is a variable which denotes the desired change in energy consumption. There are, however, two ways to view this desired change. First, it could be viewed as the desired change in *actual* energy consumption for the household. Denoting actual energy consumption at the time of the first audit as EC_{0i}^A, then in contrast to the technological relationship in (1),

$$(9) \qquad EC_{0i}^A = \bar{f}(x_{0i}, w_{0i}, h_{0i}),$$

that is, actual energy consumption would of course also depend on household characteristics, h_{0i}. With $x_{1i}^* = x_{0i} + \Delta x_i^*$ representing desired energy efficiency characteristics after the retrofits, the desired change in actual energy consumption – that is, as a result of the desired retrofits – can be written as:

$$(10) \qquad \begin{aligned} \Delta EC_i^{A*} &= \bar{f}(x_{1i}^*, w_{1i}, h_{1i}) - EC_{0i}^A \\ &= \bar{f}(x_{0i} + \Delta x_i^*, w_{1i}, h_{1i}) - \bar{f}(x_{0i,}\ w_{0i}, h_{01}), \\ &= \bar{f}(\Delta x_i^*, w_{0i}, h_{0i}) \end{aligned}$$

where the last equality in (10) follows from $w_{1i} = w_{0i}$, and $h_{1i} = h_{0i}$. Substitution from (8) now yields:

$$(11) \qquad \Delta EC_i^{A*} = \bar{f}(\tilde{g}(x_{0i}, w_{0i}, h_{0i}, D_i), w_{0i}, h_{0i}) = \bar{g}(x_{0i}, w_{0i}, h_{0i}, D_i).$$

Alternatively, the desired change in energy consumption could be viewed as the difference between calculated energy consumption – that is, based on the technological relationship used by the energy auditors – at the time of the first audit, and desired energy consumption calculated the same way but based on the desired changes in the energy efficiency properties of the house. In this case we would have:

$$\begin{aligned}\Delta EC_i^* &= f(x_{1i}^*, w_{1i}) - EC_{0i} \\ &= f(x_{0i} + \Delta x_i^*, w_{1i}) - f(x_{0i,}\ w_{0i}), \\ &= f(\Delta x_i^*, w_{0i})\end{aligned} \tag{12}$$

Substitution from (8) now yields:

$$\Delta EC_i^* \;=\; f(\tilde{g}(x_{0i}, w_{0i}, h_{0i}, D_i), w_{0i}) \;=\; \bar{\bar{g}}(x_{0i}, w_{0i}, h_{0i}, D_i)\,. \tag{13}$$

Thus, other than a possible change in functional form (which is unknown anyway), the two approaches to viewing the desired change in energy consumption are observationally equivalent.

Now the calculated change in energy consumption between the two audits, $\Delta EC_i = EC_{1i} - EC_{0i}$, is only non-zero (and in fact negative) if the i$^{\text{th}}$ house actually undertakes some retrofitting, and this only occurs if the desired change (ΔEC_i^{A*} or ΔEC_i^*) is negative and sufficiently large (say less than or equal to $-\theta$, where $\theta > 0$), so that $D_i = 1$. Thus, focusing on (13) rather than (11), in view of the way the EGH energy consumption data are constructed, we have:

$$\begin{aligned}&\Delta EC_i^* \;=\; \bar{\bar{g}}(x_{0i}, w_{0i}, h_{0i}, D_i) + e_i \\ &\text{where:} \\ &\Delta EC_i = \begin{cases}\Delta EC_i^* \text{ if } \Delta EC_i^* \le -\theta,\ \mathit{i.e., if}\ D_i = 1 \\ \\ 0 \text{ otherwise}\end{cases}\end{aligned} \tag{14}$$

Thus we have now replaced (7) and (8) with an alternative Tobit model in which the estimated parameters will reveal the effect of existing house characteristics as well as of homeowner characteristics on the amount of energy savings as a result of retrofitting, that is, on energy efficiency upgrade decisions. With this specification, the effects of the changes that were made to all the different energy efficiency characteristics of the house as a result of the retrofitting are aggregated in a useful way, thereby avoiding the need for estimation of a simultaneous system. Although ΔEC_i^* is unobserved, the data that are needed to estimate (14) include only the calculated change in energy consumption between the two audits (which equals zero for those that did not undertake retrofits and complete the second audit), and information on characteristics of the house and household occupants at the time of the first audit. Note that while estimation of an equation such as $\Delta EC_i = f(\Delta x_{i,} \; w_i)$ would be tantamount to estimating the technological relationship used to calculate energy consumption in the EGH data, such is not the case with the specification in (14) where the *change* in calculated energy consumption depends on the values of the variables at the time of the first audit, rather than on the change in the values of these variables between the two audits.

A remaining issue with the specification in (14) concerns the endogenous binary variable, D_i. Since D_i will be equal to unity for all observations where $\Delta EC_i \neq 0$, it cannot be identified separately from a standard intercept. However, since the variable being explained is the change in energy consumption, an intercept would not normally appear in the function $\overline{\overline{g}}(.)$. In this case, D_i can be included in (14), and its coefficient will indicate the effect of the EGH program on the change in energy consumption.

Finally, an alternative to estimation of the full Tobit model as specified in (14) is to restrict analysis to those households that participated in the second audit ($D_i = 1$). In this case the model becomes

$$(15) \qquad \Delta EC_i \;=\; \overline{\overline{g}}(x_{0i}, w_{0i}, h_{0i}, D_i) + e_i \qquad \text{(for observations with } D_i = 1\text{)},$$

where the error term, e_i, has a truncated distribution, which can be taken into account either directly by estimating a truncated regression model, or alternatively by simply including on the right-hand side of (15) a sample selectivity correction (inverse Mill’s ratio) obtained from a first-stage estimation of a probit model of the decision to participate in the second audit. Both these approaches, along with direct estimation of (14), are considered in the empirical work that follows.

4. Results and Analysis

To proceed, it is convenient to respecify the models in (14) and (15) so that the variable of interest is energy savings (ES_i) rather than the change in energy consumption, where

$$(16) \qquad ES_i = -\Delta EC_i = -(EC_{1i} - EC_{0i}) = EC_{0i} - EC_{1i}.$$

While $ES_i = 0$ is assumed for households that did not complete the second audit (since no data on EC_{1i} are available for these households, but clearly they did not consider it worthwhile to incur the expense associated with a second audit), it would be expected that $ES_i > 0$ for houses that did complete the second audit ($D_i = 1$). In fact this proved not to always be the case. For the 32,150 (out of 174,438) households that completed the second audit, energy savings ranged from -98,398 MJ to +1,145,843.7 MJ, with mean savings of 64,272.2 MJ.[5] Further examination of the data revealed that 251 households who completed the second audit experienced either increased energy consumption (215) or no change in energy consumption (36).

It is puzzling why some households would choose to undertake a costly second energy audit if their energy consumption had not decreased, although presumably they were unaware that the retrofits that they made had not reduced their energy consumption. One possibility is that associated with their energy retrofits, they also undertook renovations of their house, and in fact increased its size. Indeed, it is possible that the size-increasing renovations were the main objective, and that energy saving was a secondary motive, that may or may not have been enhanced by the existence of the EGH program. An investigation of this issue reveals that floor area decreased for 621 households (1.9% of the houses that completed the second audit) by an average of 10.1 square metres (with the decrease ranging between 0.1 and 155 square metres), while it increased for 1,041 households (3.2%) by an average of 29.1 square metres (ranging from 0.1 to 876.4 square metres). However, these changes in floor area occurred in very few of the houses where energy consumption increased. Not too surprisingly, for all houses where the floor area decreased, energy use also decreased. Of the 1,041 houses that experienced an

---

[5] For reasons associated with inconsistency of the values of some of the variables in the data set (especially those used to determine floor area) the sample was reduced to 170,438 households, of which 32,150 (18.9%) undertook the second audit.

increase in floor area, energy use decreased in 997 of them and increased in only 44. Thus, changes in floor area did not occur in the majority of houses in which energy use did not decrease between the two audits.

Another possible explanation for why some houses experienced an energy consumption increase between the two audits is that their participation in the EGH program occurred early, before the incentives were provided, and consequently they completed the two audits just for information purposes. Although information concerning dates in the EGH database refers to recording in the database rather than the dates of the audits, this reveals that just over half of these 251 houses appear to have completed their first audit prior to Fall 2003 when the incentives were introduced, but less than 30% had completed the second audit by this time. Thus, this explanation also does not fully account for the observed energy consumption increases.

In view of the relatively small number of households that experienced increased energy consumption between the first and second audit (less than 0.8% of those with two energy audits), they were excluded from subsequent analysis. The distribution of energy savings for the remaining 31,899 households, which exhibits a long right tail – reflecting the relatively few houses that attained very large energy savings – is shown in Figure 3.

Due to the floor area changes for some households, and the fact that energy consumption, and hence potential energy savings through retrofits, is likely to depend on floor area – since the main residential use of energy is for space-heating purposes – it is useful to calculate energy intensities, which are measured as energy consumption (MJ) per square metre. These values are displayed in Table 5 for all households, households that completed only the first audit, and for houses that completed both audits, both before and after they undertook any retrofits. As well as total energy intensity, space-heating energy intensities are also calculated. Both types of energy intensities are higher on average at the time of the first audit for households that undertook both audits than for households that completed just the first audit, and both measures decrease between audits for households that retrofitted their house and completed the second audit. Interestingly, even though the floor area decreased in some houses, energy intensities decreased in all houses that retrofitted.

Figure 3: Distribution of Energy Savings

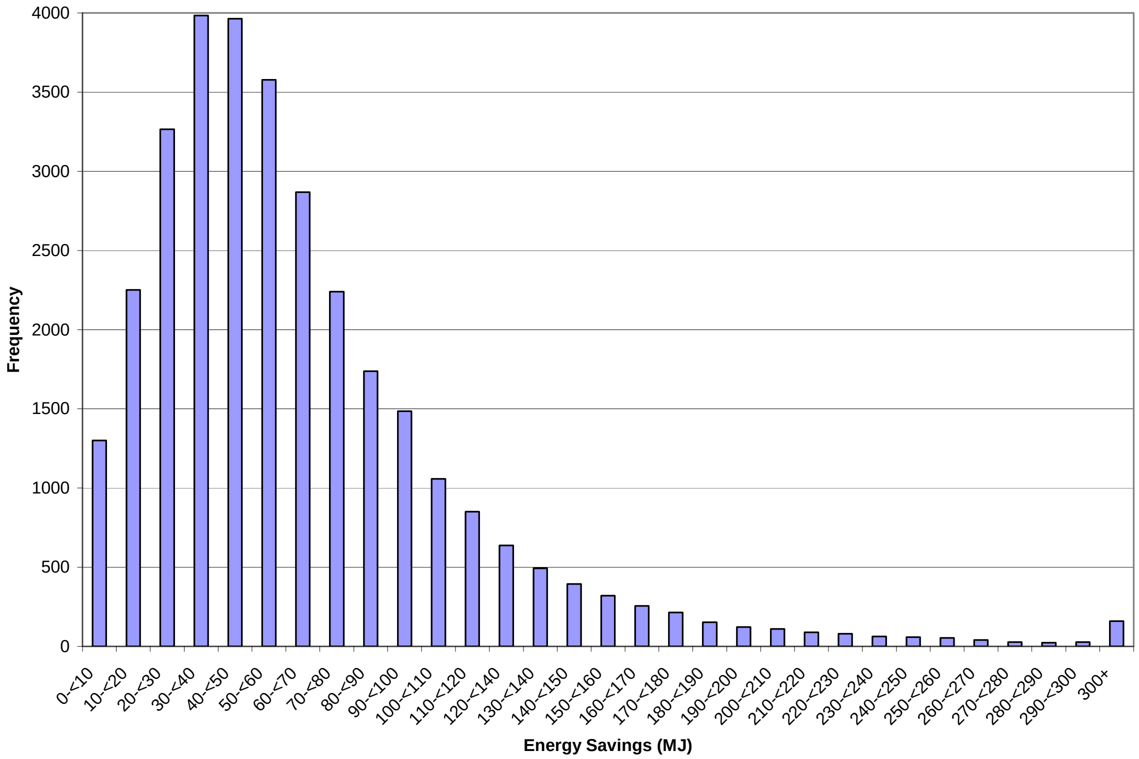


Table 5: Energy Intensities

Intensity Type	Subset of Houses	Mean	Minimum	Maximum
Energy Intensity	*1st Audit*			
	– All houses	1058.2	141.1	6767.2
	– Houses with only 1st audit	1037.4	141.1	6767.2
	– Houses with both audits	1148.3	241.0	6418.6
	2nd Audit			
	– Houses with both audits	828.2	132.8	3200.1
Space Heating Energy Intensity	*1st Audit*			
	– All houses	758.6	56.7	5952.8
	– Houses with only 1st audit	739.2	56.7	5952.8
	– Houses with both audits	842.5	117.7	5770.6
	2nd Audit			
	– Houses with both audits	526.0	56.3	2546.6

Turning to the specification of the model to be estimated, as defined in (14) or (15), it is necessary to consider the particular variables to include in the function $\overline{\overline{g}}(x_{0i}, w_{0i}, h_{0i}, D_i)$, where x_{0i} refers to characteristics (energy-using properties) of the house, w_{0i} refers to other factors that may affect energy consumption, h_{0i} refers to characteristics of the household, D_i is the dummy variable that is equal to 1 if the household undertook retrofits, and the subscript "0" indicates that all variables are defined at the time of the first audit. While x_{0i} and w_{0i} are variables contained in the EGH dataset, information on household characteristics is not collected during the audits. Consequently, information on the variables contained in h_{0i} must be obtained from elsewhere. Specifically, information from the 2001 Canadian census is used, being matched to the data in the EGH dataset according to the Forward Sorting Area (FSA), which comprises the first three characters of the six-character Canadian postal code. This is the most detailed level at which particular houses in the EGH dataset can be identified. This matching process means that specific details of the households undertaking the audits cannot be determined, but rather average household information for households living in that area is utilized. Summary statistics for the variables included in the analysis are presented in Table 6.

Table 6: Details of Variables Included in the Analysis

Type	Variable	Description	Mean	Minimum	Maximum
	ES	Amount of energy savings (GJ)	12.1516	0	1145.84
D_i	RETROFIT	=1 if made energy-saving retrofits	0.1874	0	1
x_{0i}	FLOOR	Floor area (sq. m.)	221.69	20.3	1799.6
	FLOOR2	Squared floor area	56361	412.1	3238560
	DWAGE	Age of dwelling	44.76	0	305
	DWAGE2	Age of dwelling squared	2752.61	0	93025
	FUREFF1	Furnace Efficiency at 1st audit (%)	80.84756	0	100
	DHWEF1	DWH efficiency at 1st audit (proportion)	0.609682	0	1
	CEILINS1	Average ceiling insulation at 1st audit (RSI)	4.04793	0	21.22
	BWINS1	Average basement wall insulation at 1st audit (RSI)	1.500011	0	391
	MWINS1	Average main wall insulation at 1st audit (RSI)	1.823022	0	8.81
	CONS1	Energy consumption at 1st audit (GJ)	219.77	42.47	1426.02
	CONS12	Energy consumption at 1st audit squared	55800	1803.8	2033542
	INT1	Energy intensity at 1st audit (GJ/sq. m.)	1.06	0.1411	6.77
	INT12	Energy intensity at 1st audit squared	1.30	0.0199	45.8
w_{0i}	OCCUPS	Number of occupants	3.41	1	9
	TEMP	Main floor thermostat set point (degrees Celsius)	20.91	1.1	28.9
	PROV_PEI	Located in Prince Edward Island	0.0029	0	1
	PROV_NF	Located in Newfoundland	0.0112	0	1
	PROV_NS	Located in Nova Scotia	0.0291	0	1
	PROV_NB	Located in New Brunswick	0.0099	0	1

	PROV_ON	Located in Ontario	0.3289	0	1
	PROV_MB	Located in Manitoba	0.0703	0	1
	PROV_SK	Located in Saskatchewan	0.0776	0	1
	PROV_AB	Located in Alberta	0.1980	0	1
	PROV_BC	Located in British Columbia	0.1506	0	1
h_{0i}	PLT5	% of FSA aged under 5 years	0.0559	0.0127	0.1257
	P20TO64	% of FSA aged 20 to 64 years	0.6095	0.3221	0.8670
	P65UP	% of FSA aged 65 years and older	0.1317	0	0.6242
	AVINC	Average household income ($000)	63.11	23.36	279.76
	AVINC2	Average household income squared	4383.0	545.88	78265.1
	TRCERT	% of FSA with trade certificate	0.1192	0	0.3144
	COLL	% of FSA with college education	0.2302	0.0714	0.4
	UNIV	% of FSA with university education	0.2694	0	0.8433
	NROOM	Average number of rooms in FSA	6.52	3	9.8
	UOILDHW	Rec. efficiency of oil-fuelled DWH	0.0015	0	0.94
	UGASDHW	Rec. efficiency of gas-fuelled DWH	0.0010	0	0.93
	UPROPDHW	Rec. efficiency of propane-fuelled DWH	0.0012	0	0.9
	UELECDHW	Rec. efficiency of elec.-fuelled DWH	0.0260	0	1
	UELECFUR	Rec. efficiency of elec.-fuelled furnace	0.8663	0	100
	UGASFUR	Rec. efficiency of gas-fuelled furnace	46.67	0	100
	UOILFUR	Rec. efficiency of oil-fuelled furnace	2.92	0	100
	UPROPFUR	Rec. efficiency of propane-fuelled furnace	0.3800	0	97.3
	UWOODFUR	Rec. efficiency of wood-fuelled furnace	0.0731	0	100
	UGR_CEIL	Rec. additional ceiling insulation (RSI × sq. m.)	179.31	0	2978.61
	BWA_11	Basement wall area - if upgrade ≤ 1.9 RSI (sq. m.)	21.47	0	92.85
	BWA_19	Basement wall area - if upgrade > 1.9 RSI (sq. m.)	0.7491	0	82.55
	MWA_4	Main wall area - if 4-inch thick walls (sq. m.)	4.41	0	67.84
	MWA_6	Main wall area - if 6-inch thick walls (sq. m.)	1.87	0	57.08
	COSTSAVE	Predicted cost savings (if follow all recs.)	850.39	-3899.64	14329.13
	DAREA	Change in floor area between audits (sq. m.)	0.1130	-155	347.4
	DAREA2	Change in floor area between audits squared	15.74	0	120686.8
	DVAREAP	=1 if floor area increased between audits	0.0059	0	1
	DVAREAN	=1 if floor area decreased between audits	0.0036	0	1

Note: “Rec.” means upgrade recommendation; DWH is domestic water heating; “gas” refers to natural gas.

Noting that the selected variables may play a role either in the decision to retrofit ($D_i = 1$) or the amount of energy savings (ΔES_i), the first group of variables listed in Table 6, x_{0i}, which refer to the energy-using properties of the house, are based on data collected at the first energy audit. These include the floor area and floor area squared, since energy use is likely to increase with floor area, but not necessarily at a constant rate, as well as the age of the dwelling and its square, again since older dwellings – if not retrofitted – are likely to use more energy, and again this increase may not be constant with increasing age. The furnace efficiency and the efficiency of the water heater, as well as average ceiling, foundation (basement) wall, and main wall insulation, all of which reflect the current energy-using state of the house and hence reflect the need for energy-saving retrofits, are also included. While there is a wealth of other detailed information on the house, we have combined these effects here by including total energy consumption at the first audit, as well as its square. Since total energy consumption is calculated using a technological relationship, this effectively aggregates the roles of these other characteristics. A larger value of initial energy consumption might mean an increased likelihood of retrofitting, as well as an increase in the amount of energy savings that could be achieved through retrofitting, but this relationship is complicated by differences in floor area, since the possibilities may be quite different for a large house and a small house that both have high initial energy consumption. To allow for this possibility, both energy intensity – that is, energy used per square metre – as well as its square are included along with energy use and floor area. It would be expected that a higher level of initial energy intensity would increase both the likelihood of retrofitting and the energy savings that could be achieved, and these effects might be expected to be nonlinear.

The second group of variables in Table 6, w_{0i}, refer to other factors that may affect energy consumption, again based on data collected at the first audit. Here, we include the number of occupants of the household, on the expectation that more occupants would increase energy use, as well as the main floor thermostat set point, since a higher setting would typically result in greater amounts of energy being used for space heating, and hence the potential savings from energy-efficient retrofitting of the house. To capture weather effects, with lower temperatures increasing the need for space heating, we include dummy variables for each province (Quebec and the Territories are omitted). Of course these variables capture more than just weather differences between the provinces, and likely reflect cultural factors as well as possibly different costs associated with retrofitting in different regions.

The final group of variables, h_{0i}, are designed to reflect characteristics of the household, although more generally these also capture other factors that would affect the probability of retrofitting and the amount of retrofitting that is done, and hence the energy savings that are likely to be achieved. Variables included here include the age distribution in the same FSA – particularly the proportion of children, adults, and seniors – as well as average household income and its square, higher education levels in the same FSA – represented by the proportion of residents with a trade certificate, college education, or university education – and the average number of rooms in houses in the same FSA. Additional variables included in this household category reflect some of the likely costs and even benefits involved in the energy-efficiency retrofit process. Unfortunately, since there are no direct data on costs and prices included in the EGH dataset, it is necessary to infer these indirectly. Specifically, we include a number of variables that indicate the extent of the upgrade recommendations made by the energy auditor, since an increase in these variables would typically be associated with a higher cost of upgrading, although also – at least in some cases – a greater benefit from making the upgrades.[6]

In this category of upgrade recommendation variables we include the recommended efficiency of particular upgrades to domestic water heaters, which are fuel-specific (oil, natural gas, electricity, and propane), as well as the recommended efficiency of fuel-specific replacement space-heating furnaces, including wood-based heaters.[7] The amount of the recommended upgrade to ceiling insulation, calculated as the product of the area to be insulated and the recommended increase in insulated value (RSI), is also included, along with variables pertaining to main-wall and basement-wall insulation. These latter variables are calculated separately depending on the thickness of the inside walls (typically foundations are poured concrete, and insulated walls are built inside these), since this affects the amount of insulation that can be installed, as well as the amount of the recommended upgrade to the insulation.[8] To capture the costs of various other upgrade recommendations that may have been made by the energy auditor, the predicted cost savings if all upgrade recommendations are followed (a variable which is provided to homeowners who complete the first audit, and is therefore included in the EGH database) is also included. Finally, in view of the earlier discussion about a number of homeowners changing the floor area of their home, variables that indicate the change in floor

---

[6] With this approach the actual dollar cost per unit is absorbed into the coefficients on these variables. An alternative approach is to attempt to estimate the costs of each type of retrofit in each location, although this might be expected to have an extremely high degree of inaccuracy. This procedure was used by Guler et al. (1999), whose estimates were subsequently adopted by Aydinalp et al. (2001).

[7] Note that not all fuels are available in all areas, the most notable being natural gas which is not available in the Atlantic Provinces.

[8] These variables could of course be considered to be part of the set of variables representing house characteristics rather than household characteristics, but in the exposition here we have focused on their roles in terms of affecting the cost of retrofitting, since they are based on upgrade recommendations.

area and its square, again to allow for nonlinear effects, are also included. This change in floor area is viewed here as a realization of the homeowner's desire to increase their floor area, which is assumed to have been determined prior to making any energy-efficiency retrofits. As an alternative to the actual change in floor area, a dummy variable that indicates if the floor area was increased, and another dummy variable that indicates if the floor area was decreased was also considered.

4.1 Probability of Retrofitting (and Completing the Second Audit)

Since it forms part of the truncated regression model, and its parameters are used in the determination of the sample selectivity correction using the two-step procedure when (15) is estimated, we first examine the results obtained with the model of the probability that a household will retrofit, as evidenced by their completion of the second audit.[9] Thus the specification for this model is:

$$
\begin{aligned}
& D_i^* \;=\; g^*(x_{0i}, w_{0i}, h_{0i}) + e_i \\
(17) \quad & \text{where:} \\
& D_i = \begin{cases} 1 \text{ if } \; D_i^* > 0, \; \text{i.e., if } ES_i > 0 \\ \\ 0 \;\; \text{otherwise} \end{cases}
\end{aligned}
$$

and e_i is a random error term assumed to be normally distributed. Although in principle all the explanatory variables in Table 6 could have a potential role in explaining the decision to retrofit, save energy, and undertake the second audit, there are some operational considerations that limit the types of variables that can be included in the estimation. In particular, any variable that is non-zero only for households that retrofit – even if not necessarily for all households that retrofit – will lead to a problem with quasi sample separation. In the context of the variables in Table 6, this will occur if the dummy variables DVAREAP or DVAREAN are included, since they only take non-zero values for (some of the) households that made retrofits (since the change in area is not observed, and is assumed to be zero, for households that did not complete the second audit). As a result, whenever these variables are non-zero, the estimation procedure will match the

---

[9] Gamtessa and Ryan (2007) contains an earlier empirical analysis using this type of model with the EGH data.

observation with $D_i = 1$, regardless of the coefficient on these variables. Consequently, the coefficients on these variables are undetermined and estimation with them included will not converge. This is disappointing, and possibly leads to a model misspecification, since the decision to undertake energy-saving retrofits may be enhanced by the decision to renovate. There are various alternative solutions to this problem, although all of them are ad hoc in nature.[10] Fortunately, this problem can be avoided here by the use of the DAREA variable, since this takes positive, negative, and zero values for households that retrofitted, although it is always zero for households that did not complete the second audit. Unfortunately, DAREA2, the squared value of DAREA cannot be included since its positive values only occur for households that made retrofits, again leading to a quasi sample separation problem.

The results obtained when (17) is estimated as a probit model are presented in Table 7. Rather than the parameter estimates, which have no direct interpretation other than their signs, Table 7 contains the marginal effects, that is, the effect on the probability of retrofitting of a marginal change in each of the continuous explanatory variables, or a change from zero to one for the dummy variables.

---

[10] Zorn (2005) reviews these alternatives and suggests instead the use of a penalized likelihood function, which he shows may avoid the misspecification problem and allow estimates to be obtained of the parameters on the variables that cause the quasi sample separation. The use of this approach will be considered in future research.

Table 7: Marginal Effects from Probit Estimation of (17)

Variable	Coefficient	Standard Error	Asymptotic t ratio	Prob Value
Constant	-0.3880	0.0791	-4.907	0
DWAGE	0.0008	0.0001	5.472	0
DWAGE2	-0.000006	0.0000009	-6.883	0
FUREFF1	-0.0058	0.0001	-41.113	0
DHWEF1	0.1061	0.0098	10.835	0
CEILINS1	-0.0017	0.0007	-2.226	0.026
BWINS1	-0.0029	0.0003	-10.482	0
MWINS1	-0.0218	0.0024	-9.244	0
CONS1	0.0003	0.00004	7.576	0
CONS12	-0.0000007	0.0000001	-10.636	0
INT1	0.0650	0.0103	6.305	0
INT12	-0.0293	0.0031	-9.573	0
OCCUPS	-0.0118	0.0009	-13.673	0
TEMP	0.0073	0.0017	4.298	0
PROV_PEI	-0.0713	0.0164	-4.34	0
PROV_NF	-0.0107	0.0121	-0.885	0.3762
PROV_NS	0.0489	0.0091	5.349	0
PROV_NB	0.1117	0.0133	8.378	0
PROV_ON	0.0829	0.0050	16.648	0
PROV_MB	0.1296	0.0074	17.592	0
PROV_SK	0.1099	0.0081	13.499	0
PROV_AB	0.0703	0.0061	11.47	0
PROV_BC	0.0729	0.0062	11.836	0
PLT5	-0.7125	0.1682	-4.235	0
P20TO64	0.2704	0.0661	4.09	0
P65UP	0.3217	0.0613	5.249	0
AVINC	-0.0011	0.0003	-3.792	0.0001
AVINC2	0.000001	0.000001	0.976	0.3292
TRCERT	0.0464	0.0644	0.721	0.4709
COLL	0.0420	0.0339	1.24	0.2149
UNIV	-0.0375	0.0206	-1.826	0.0679
NROOM	0.0307	0.0030	10.226	0
UOILDHW	0.0869	0.0318	2.727	0.0064
UGASDHW	-0.0682	0.0107	-6.384	0
UPROPDHW	-0.0518	0.0373	-1.389	0.165
UELECDHW	-0.0163	0.0072	-2.279	0.0227
UELECFUR	-0.0007	0.0001	-6.308	0
UGASFUR	0.0006	0.00003	20.078	0
UOILFUR	-0.0005	0.0001	-7.297	0
UPROPFUR	-0.0003	0.0002	-1.663	0.0964
UWOODFUR	-0.0033	0.0005	-6.185	0
UGR_CEIL	-0.000004	0.000006	-0.694	0.4875
BWA_11	-0.0005	0.0001	-8.334	0
BWA_19	-0.0004	0.0002	-2.275	0.0229
MWA_4	-0.0019	0.0002	-11.587	0
MWA_6	-0.0011	0.0002	-5.77	0
COSTSAVE	0.00003	0.000002	14.978	0
DAREA	0.0042	0.0003	16.505	0

As can be seen from the column in Table 7 labelled "prob value", the only variables that are not significant at a 10% level or better are the squared value of average income, the proportion of households with a trade certificate, or with college education, houses located in Newfoundland, the recommended upgrade to ceiling insulation, and the recommended upgrade efficiency measure for propane water heaters. In addition, floor area and squared floor area were insignificant and were omitted, as their effect appears through the energy intensity variables, INT1 and INT12. Most other variables are significant at a 1% level, although the proportion with university education and the recommended upgrade efficiency measure for propane furnaces are only significant at 10%, three other variables are only significant at a 5% level. In terms of overall significance, the maximized value of the log likelihood function is –76409.11, and a joint test of the significance of all the parameters yields a Likelihood Ratio statistic of 11406 with 47 degrees of freedom, which is significant at better than a 1% level. The model correctly predicts no retrofitting for all but 251 of the households that did not retrofit, but only predicts retrofitting for 335 of the 31,899 that did retrofit, resulting in an overall prediction success rate of 81.31%.

In terms of the estimated marginal effects in Table 7, the findings are for the most part consistent with expectations. The probability of retrofitting increases with the dwelling age, total energy consumption at the first audit, and energy intensity at the first audit, but all at a decreasing rate. This probability decreases with higher initial levels of furnace efficiency, and ceiling and wall insulation, but increases with higher initial water heater efficiency. A higher temperature set point increases the probability of retrofitting, but this probability decreases with the number of occupants. There are obvious age distribution effects, with the probability increasing with higher proportions of adults and seniors, but decreasing with a larger proportion of children. A surprise perhaps is the negative sign on average income, indicating that higher incomes decrease the probability of retrofitting. This might indicate that the desire for energy savings is greater among those who are less able to afford increasing energy costs, but this may seem counter-intuitive in view of the capital costs that must be incurred to retrofit, unless it is believed that these will be offset by the energy savings that will eventuate because of the retrofits. Some additional support for this interpretation is provided by the positive sign on the cost savings variable. Also, the probability of retrofitting decreases with the proportion of individuals that have university education. Relative to Quebec, which has a greater reliance on electricity and has relatively low electricity costs, the probability of retrofitting is significantly higher in all other provinces except Prince Edward Island, where it is lower, and in Newfoundland where it is not significantly different. In terms of upgrade recommendations, the probability increases with higher upgrade efficiency recommendations for gas furnaces and for oil-based water-heating systems, but decreases with higher upgrade efficiency recommendations for other types of furnaces and for electric or gas water heating systems. Larger wall areas in any part of the house, and apparently regardless of their current thickness, decrease the probability of retrofitting. This may reflect the possibly higher cost of retrofitting in these cases, or possibly a realization of the limited value of other types of retrofitting if the walls need to be re-insulated but that process is viewed as too disruptive or costly. Finally, proposed renovations that will increase the floor area significantly increase the probability of making energy-saving retrofits.

4.2 Explaining Energy Savings

We now turn to the results obtained using the models that explain energy savings, a variable which effectively aggregates the effects of all the energy-saving retrofits that were undertaken. As noted in Section 3, several different estimation methods were used. In addition to direct estimation of the Tobit model, as specified in (14), we also consider estimation of (15) using both direct estimation of a truncated regression model, and the more common estimation of (15) with a sample selectivity correction (Inverse Mills ratio), based on the parameters of the estimated probit model reported above, included on the right-hand side as an additional explanatory variable. The estimated marginal effects obtained from these estimations, which show the effect of each explanatory variable on the expected energy savings, are presented in Table 8. For the estimation using the sample selectivity correction, both the direct marginal effects and the total marginal effects are shown, where the total effect includes both the direct effect of a change in a variable on energy savings as well as the indirect effect that arises since a change in a variable that is also included in the first-stage probit model affects the sample selectivity term, which in turn affects energy savings. For comparison, the estimated marginal effects from the probit model, as reported in Table 7, are included in the final column of Table 8.

Table 8: Marginal Effects from Tobit, Truncation, and Sample Selection Estimations

Variable	Tobit	Truncation	Sample Selection		Probit from Table 7
			direct	total	
Constant	-27.9524**	-68.4160**	-93.8896**		-0.3880**
FLOOR	-0.0091	0.0689**	-0.0503**		
FLOOR2	0.0000	-0.0001**	-0.00003†		
DWAGE	0.0353**	-0.1588**	-0.0599†	-0.1585**	0.0008**
DWAGE2	-0.0004**	0.0004*	-0.0004	0.0004	-0.000006**
FUREFF1	-0.4132**	-0.7387**	-1.3518**	-0.6082**	-0.0058**
DHWEF1	8.2132**	17.1686**	30.9706**	17.4028**	0.1061**
CEILINS1	-0.0579	0.6871**	0.4559**	0.6687**	-0.0017*
BWINS1	-0.1844**	-0.1811*	-0.5313**	-0.1595	-0.0029**
MWINS1	-1.0655**	3.9648**	1.1475	3.9409**	-0.0218**
CONS1	0.0422**	0.1924**	0.2377**	0.1955**	0.0003**
CONS12	-0.00003**	0.0001**	0.0002**	0.0002	-0.0000007**
INT1	0.5342	27.5840**	0.2279	-8.0877†	0.0650**
INT12	-0.6808*	-0.5397	3.1278**	6.8696**	-0.0293**
OCCUPS	-0.7014**	0.9127**	-0.6985*	0.8121†	-0.0118**
TEMP	0.4372**	0.5089	1.3946**	0.4654	0.0073**
PROV_PEI	-7.0080**	-12.3681*	-23.2309**	-12.3836†	-0.0713**
PROV_NF	-1.2603	-8.0352**	-8.5830**	-7.1848*	-0.0107
PROV_NS	3.1639**	-1.6805	3.5904	-2.1959	0.0489**
PROV_NB	6.3880**	12.7029**	23.0600**	10.8649*	0.1117**
PROV_ON	5.1667**	-4.0615**	5.9818*	-4.1161	0.0829**
PROV_MB	6.7875**	-8.4124**	5.8132†	-8.3124†	0.1296**
PROV_SK	6.0329**	-9.9889**	2.8244	-9.4126*	0.1099**
PROV_AB	4.0616**	-6.7097**	2.4101	-5.9639*	0.0703**
PROV_BC	4.5798**	4.3470**	11.2564**	2.6890	0.0729**
PLT5	-44.0775**	-61.6315†	-132.7582**	-41.6229	-0.7125**
P20TO64	19.2048**	28.7535*	62.0903**	27.5023	0.2704**
P65UP	23.2735**	36.1705**	74.3314**	33.1844†	0.3217**
AVINC	-0.0834**	-0.2403**	-0.3431**	-0.2064*	-0.0011**
AVINC2	0.0001	0.0006*	0.0006*	0.0005	0.000001
TRCERT	1.5804	-9.0026	-7.2047	-13.1446	0.0464
COLL	3.7436†	-11.4027†	-1.7631	-7.1342	0.0420
UNIV	-2.5455†	-10.1331*	-15.3506**	-10.5503*	-0.0375†
NROOM	2.2695**	5.2124**	8.6464**	4.7257**	0.0307**
UOILDHW	5.1271*	-16.1746*	-6.7953	-17.9058*	0.0869**
UGASDHW	-5.3555**	-21.4967**	-30.6021**	-21.8761**	-0.0682**
UPROPDHW	-3.8903	15.8840*	7.1145	13.7393	-0.0518
UELECDHW	-1.2659**	0.9087	-1.4113	0.6736	-0.0163*
UELECFUR	-0.0353**	0.2117**	0.1321**	0.2184**	-0.0007**
UGASFUR	0.0356**	0.0407**	0.0878**	0.0122	0.0006**
UOILFUR	-0.0420**	-0.1172**	-0.1866**	-0.1231**	-0.0005**
UPROPFUR	-0.0212†	-0.0592†	-0.0962*	-0.0572	-0.0003†
UWOODFUR	-0.2365**	0.2791*	-0.1387	0.2831	-0.0033**
UGR_CEIL	0.0003	0.0106**	0.0089**	0.0095	-0.000004
BWA_11	-0.0277**	0.0316**	-0.0580**	0.0053	-0.0005**
BWA_19	-0.0262*	0.0383	-0.0316	0.0254	-0.0004*
MWA_4	-0.1231**	-0.1469**	-0.3897**	-0.1467†	-0.0019**
MWA_6	-0.0729**	-0.1200**	-0.2800**	-0.1375*	-0.0011**
COSTSAVE	0.0032**	0.0186**	0.0248**	0.0205	0.00003**
DAREA	0.1951**	0.0065	0.3950**	-0.1445	0.0042**
Log Likelihood	-238,452.9	-152,616.8	Probit (using the same variables):-76,406.45		

Note: **, *, and † denote significance at the 1%, 5%, and 10% levels, respectively.

The estimated marginal effects in Table 8 reveal some large differences between the various estimation methods for some of the variables. To some extent this may reflect the effects of the restrictions inherent in the Tobit model, which ensure that the coefficient on each variable is the same in the determination of the amount of energy savings as well as in the decision of whether or not to save any energy by making retrofits. While, at least in terms of sign, this restriction appears appropriate in many cases – for example, a higher initial level of energy consumption would be expected to increase the likelihood of making retrofits as well as the amount of retrofitting that is done if any retrofits are made, this may not be the case for all variables, and the magnitudes and significance of the effects may differ in the two decision-making processes. This difference in the roles of variables in the two components of the model was the motivation for the Cragg (1971) model, which essentially replaces the Tobit model with the truncated and probit models. A Likelihood Ratio (LR) test of these restrictions can be obtained (Greene, 2003) by summing the log likelihoods for the probit and truncated models and comparing this total to the log likelihood for the restricted (Tobit) model. The last row of Table 8 contains the three log likelihood function values, where the value reported for the probit model refers to the case where the same variables are included in the probit specification as in the truncated and Tobit models. These yield a LR statistic of 26,841.3 with 50 degrees of freedom, which is significant at a 1% level, indicating that the restrictive Tobit formulation is not appropriate.

The estimated marginal effects from the truncated model and from the sample selection model (total effects) are remarkably similar for most of the variables. Although there are some differences in the levels of significance of the effects, with many more marginal effects being significant, or significant at a higher level, in the truncated model, the only major differences occur for INT1 which is positive and significant at the 1% level in the truncated model but negative and only significant at the 10% level in the selection model, and INT12 which is positive and significant at the 1% level in the selection model but negative and insignificant in the truncated model.

Focusing on the results of the truncation model, energy savings are seen to increase for larger houses, but at a decreasing rate, so that energy savings start to decrease for houses larger than approximately 345 square metres. Energy savings decrease as the age of a building increases, which seems somewhat surprising, as older buildings – unless previously retrofitted – would be expected to have greater opportunities for energy savings. However, the positive sign on DWAGE2 indicates that this negative effect becomes smaller for older buildings. Higher furnace efficiency at the time of the first audit is associated with lower energy savings, although the opposite effect is found for water heater efficiency. Higher initial levels of ceiling insulation or main wall insulation also increase energy savings, but lower initial basement wall insulation levels result in higher energy savings. In terms of energy consumption and intensity, houses with higher initial levels of energy consumption achieve greater energy savings, and this effect increases at an increasing rate, while the most energy-inefficient houses achieve the greatest energy savings.

Compared to Quebec, houses in New Brunswick or British Columbia experience greater energy savings – holding all other factors constant – while those in other provinces, except for Nova Scotia experience significantly lower energy savings. Houses with more occupants experience greater energy savings, but the thermostat set point plays no significant role.

Turning to the household variables, significantly greater energy savings are associated with households having a larger percentage of seniors and occupants aged 20 to 64. Perhaps somewhat surprisingly, energy savings decrease as average income increases, although this effect decreases in magnitude as incomes increase. Higher education is generally associated with lower energy savings, while houses with more rooms tend to experience larger energy savings. In terms of the recommended upgrades, higher levels of many of these are associated with significantly larger energy savings, as might be expected, but some have the opposite effect, such as higher recommended efficiency of oil or natural gas water heaters, oil or propane furnaces, and higher recommended insulation of main walls, a type of retrofit that – as noted earlier – in contrast to many other types of retrofit is likely to be very disruptive to household operations. Finally, higher predicted energy cost savings if all upgrade recommendations are followed, which would generally be associated with houses that are very energy inefficient, has a significant positive effect on energy savings, holding all other factors constant.

5. Summary and Conclusions

In this paper we utilize data collected in two rounds of residential energy audits under the Canadian EGH Program to model and empirically investigate the determinants of the decision to retrofit for those that completed the first audit, as well as the factors that affect the energy savings that were achieved by homeowners who participated in the program and completed the second audit. This analysis is hampered by the fact that the energy consumption data in the EGH dataset are inferred from technological relationships rather than reflecting actual measures. Consequently, estimation of a model where energy consumption is specified to be a function of the energy-efficiency characteristics of the house along with behavioural factors, and possibly other measures, will simply result in an imprecise estimate of the technical relationship that was used to infer the energy consumption values, and behavioural factors will appear to have no role to play. A similar result will occur with estimation of a model of energy savings, calculated as the difference between energy consumption at the second audit and at the first audit, where the explanatory variables include differences in energy-efficiency characteristics of the house as well as behavioural and other factors. An additional problem with this latter specification is that many of the explanatory variables pertaining to changes in various energy-efficiency characteristics will be endogenous, since they result from decisions made by the household that will depend on a variety of factors.

To deal with these problems we derive a model where inferred energy savings depend on energy-efficiency characteristics of the house at the time of the first audit, as well as household variables and factors reflecting aspects of the expected cost of retrofits. Unfortunately, the EGH dataset does not contain information on household characteristics, so that in order to include these types of variables in the empirical analysis it is necessary to supplement the EGH data with average information about households obtained from the Canadian census and matched to observations in the EGH dataset using part of the postal code for each house. Since energy savings are only observed for houses where the second audit was completed, the resulting model has a Tobit (limited dependent variable) specification. A drawback of this type of model is that the explanatory factors have the same directional effect on the probability of undertaking retrofits and the amount of retrofitting that is undertaken for those houses where any retrofits are made.

These restrictions can be tested by separately estimating a probit model and a truncated regression model as well as the Tobit model, and a test of these restrictions rejects the more restrictive Tobit formulation. An alternative is to focus on the truncated regression results, or to use a sample selectivity model where a correction, derived from the probit model of whether or not households undertake any retrofits, is included as an additional explanatory variable in the energy savings equation that is estimated only for those with positive values of energy savings. The two models yield similar marginal effects, although the significance of these effects differs between the two formulations.

Generally the results show that factors other than just the energy-efficiency characteristics of the house prior to the retrofits have a significant effect on energy savings, including some of the household or socio-economic factors as well as variables that reflect the expected cost of the retrofits and the energy-cost savings that might be achieved. Thus, even though the energy consumption information at each audit is inferred from a technological relationship that only involves variables that pertain to the energy-efficiency characteristics of the house, there appears to be scope to use the data with an appropriately-specified model to ascertain the role of various other factors that may play a role in determining the energy savings that are achieved through residential retrofitting.

As noted at various points in our analysis, there are a number of issues that we have not considered but which would be useful to incorporate into our analysis. First, selection into the first energy audit is non-random, and this has not yet been taken into account. Unfortunately we do not have corresponding data for houses that did not complete this first audit, so alternative ways to deal with this issue need to be considered. Second, information on household characteristics is based solely on averages obtained from the Canadian census, so this severely limits the strength of any conclusions that can be drawn regarding these variables. Since the EGH dataset does not contain this type of information, there is no obvious alternative to this procedure, but it may be possible to fine tune it to some extent, and this issue is certainly worth investigating. Third, the sheer size of the EGH dataset has implications for the statistical significance of the estimated coefficients and various tests that are conducted. An alternative may be to subdivide the data according to geographical region or type or age of dwelling, etc., which may have an added advantage in some cases of focusing the analysis on characteristics that might be more applicable for the sub-divisions that are chosen. Finally, and perhaps most importantly, the EGH energy consumption data that are used in the analysis here are based solely on technological relationships. While the model that has been derived allows factors other than those explicitly included in these technological relationships to play a role in the determination of energy savings, it would be interesting to determine whether the general findings from this formulation hold when actual energy consumption data are used. While this obviously cannot be assessed using the EGH dataset where actual energy consumption data are not available, an interesting extension would be to use a dataset where such variables are available, to infer energy consumption using the technological relationships used by the energy auditors in the EGH program, and to compare the results obtained using the actual and inferred energy savings variables. At the very least, such an analysis might provide convincing evidence on the importance of collecting information on household as well as house characteristics as part of the energy audit process.

References

Aydinalp, M., A. Ferguson, A. Fung, and I.V. Ugursal (2001), "Energuide for Houses Database Analysis", Research Paper CREEDAC-2001-04-03, CREEDAC, Dalhousie University.

Blais, S., A. Parekh, and L. Roux (2005), "Energuide for Houses Database – An Innovative Approach to Track Residential Energy Evaluations and Measure Benefits", Paper presented at *Building Simulation 2005*, the 9th International Conference of the International Building Performance Simulation Association, Montreal, August.

Cragg, J. (1971), "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica*, 39, 829-844.

Gamtessa, S. and D.L. Ryan (2007), "Utilization of Residential Energy-Saving Retrofit Programs in Canada: Who, What, and Why", paper presented at the 9th European Meetings of the International Association for Energy Economics, Florence, Italy.

Greene, W.H., (2003), *Econometric Analysis*, 5th Edition (Upper Saddle River, NJ: Prentice Hall).

Guler, B., A.S. Fung, M. Aydinalp, and V.I. Ugursal (1999), "The Techno-Economic Analysis of Home Retrofit Activities and Associated Energy Savings in the Residential Sector in Canada", Research Paper CREEDAC-1999-12-6, CREEDAC, Dalhousie University.

Zyniewski, A. and P. Mistry (2005), "EnerGuide for Housing Preliminary Analysis", PowerPoint presentation, Natural Resources Canada.

Zorn, C. (2005), "A Solution to Separation in Binary Response Models", *Political Analysis*, 13(1): 157-170.

CBEEDAC
Department of Economics

University of Alberta

8-14 Tory Building

Edmonton, Alberta

Canada

T6G 2H4