Behavioral Models of Free Riders in DSM Programs

Michael T. Ozog, RCG/Hagler, Bailly, Inc.
Donald M. Waldman, University of Colorado

This paper presents a technique which uses statistical behavioral models to estimate the level of free riders in DSM programs. These models are a variation of the discrete-choice participation models which have been used extensively in DSM impact evaluations. These behavioral models relate the decision to undertake the conservation action promoted by the utility (whether or not it was done through the program) as a function of demographic and attitudinal variables as well as the program incentive (e.g., the rebate level or the loan buydown). The level of free ridership is then determined by finding the probability of undertaking the conservation action given that there is no program incentive.

In addition to estimating the level of free riders, these models can also be used to simulate the level of free riders when the program incentive is changed or different segments of the customer population are targeted for participation. Therefore, these models can be used for both program design and program evaluation.

These models represent a potentially powerful tool for DSM evaluation, since they quantify the level of free ridership and produce testable results. In addition, unlike other survey-based methods, these models do not require a subjective interpretation by the researchers and are not prone to hypothetical bias and cognitive dissonance. However, they may be influenced by the presence of free drivers.

The paper reviews the theory behind these models, and then presents estimated results from a recent impact evaluation of a residential audit program.

Introduction

Free riders are broadly defined as participants in a DSM program who would have undertaken some or all of the actions promoted by the program even if the program had not existed.

While free riders may not be an issue from a systems standpoint (since a decrease in energy usage is a decrease irrelevant of the source) free riders may affect the cost-effectiveness of a program. Because the participant would have undertaken the conservation action regardless of the program, the utility is subsidizing the behavior of these individuals. Therefore, from a utility perspective (i.e., the utility cost test), free riders decrease the cost-effectiveness of the program.

The most common approach to determine the level of free riders is to conduct a survey of program participants. These surveys generally rely on participants to report what they would have done without the program. As has been pointed out by Kreitler (1990) and Saxonis (1991), there are severe limitations to using surveys to measure the level of free riders. These limitations include, but are not limited to, cognitive dissonance and hypothetical bias. Cognitive dissonance occurs when respondents rationalize the decision they made, justifying that they took the correct action and would have taken it without any program inducement (even though that may not be the true situation). Thus, free ridership will be over estimated. Hypothetical bias occurs because the survey is asking a hypothetical question--what would they have done without the program--which produces a hypothetical answer. For example, the DSM program changes the relative prices, and so the participants probably cannot know what they would have done if they had been faced with relative prices that did not reflect the DSM program.

This paper presents a statistical method to determine the level of free riders in a DSM program. With this technique, the level of free riders is determined by specifying a behavioral model for taking the conservation action promoted by the DSM program. This technique involves an extension of the discrete-choice participation model which is commonly used in DSM impact evaluations to control for self-selection bias (Violette and Ozog (1989)). Unlike other approaches, this technique does not rely on participants to report what they would have done without the program, and therefore it is not subject to cognitive dissonance and hypothetical bias. However, since this model does implicitly use the action of nonparticipants to proxy the behavior of participants without the program, it may be biased by the effect of free drivers.

While other researchers (notably Train (1990) and Regional Economic Research (1991)) have developed similar discrete-choice models which can be used to develop free rider estimates, the model presented in this paper is a significant variation of these approaches. Specifically, these models use a participation variable as an independent variable in a model of the adoption of a specific conservation measure. Under this approach, the level of free ridership is found by simulating the mean probability of adopting the measure for participants when the participation variable is set to zero. While the spirit of these models is similar to that of the model presented in this paper, the model in this paper does not use a participation variable as an independent variable. A participation variable is not used because it requires the estimation a separate participation model to correct for the self-selection bias that arises when a participation variable is included as an independent variables.

This discussion begins with a review of the theoretical model used in the analysis and the specification of the key variables. The paper then shows the results produced when this technique was applied to a residential audit program.

Theoretical Model

Consider a household model of energy conservation behavior, where each household has the possibility of taking some conservation action. Households are assumed to decide whether or not to take action based on an internal benefit-cost type analysis, where the benefits include the monetary savings, while the costs include both out-of-pocket costs of the actions, as well as any time costs. The household performs the conceptual experiment of maximizing their well-being (or utility) with and without the action. Let y* measure the net utility (or net benefit) of undertaking a conservation action. That is, y is a measure of the difference in a household's "well being" between taking and not taking a conservation action. A possible model for y* is:

$$y* = g(\textit{benefits, costs, conservation ethic}) + \epsilon \qquad (1)$$

where g(.) is a function that aggregates all the determinants of utility, and ϵ is a random disturbance term.

Let x be a vector of housing and individual characteristics (including attitudes, opinions and beliefs) that influence the benefits and costs of the conservation actions. In addition to these "control" variables, an important aspect of the model that is key to distinguishing free riders from program-induced participants is the specification of the cost savings due to program participation (the "program effect"). Let z equal the reduction in the cost of adoption of the conservation action due to the DSM program. A linearized model of net benefits is then:

$$y* = \beta' x + \gamma z + \epsilon \qquad (2)$$

where (β,γ) is the vector of marginal (net) utilities, x is the individual characteristics, and z is the program effect variable. Net benefit is not observed, but what is observed is whether or not a household took action. Define the observable dichotomous variable y to be whether or not the individual takes conservation action, such that:

$$y = \begin{Bmatrix} 1 \textit{ if } y*>0 \\ 0 \textit{ if } y* \leq 0 \end{Bmatrix}. \qquad (3)$$

Thus, if the net benefit of undertaking the conservation action is positive, then the individual takes the action (y = 1); otherwise, the individual does not take the action (y = 0).

Whether or not this household takes action without a program (i.e., whether or not this household is a free rider) depends upon the evaluation of net benefit without any program inducement. This is simulated by setting the z variable equal to 0. In the model of Equation (2), this net benefit is equal to:

$$y^*_{|z=0} = \beta' x + \epsilon \qquad (4)$$

so that the probability of taking action without a program is given by the expression:

$$\begin{aligned} pr[y^*>0|z=0] &= pr[\beta' x + \epsilon > 0] \\ &= 1 - F(-\beta' X), \end{aligned} \qquad (5)$$

where F(.) is the cumulative distribution function of the random variable ϵ. This is a measure of the free-rider effect (that is, the probability of a household undertaking action without the program inducement).

Specifying Program Effects

Specifying the program effect variable z into the estimation is not straightforward. It is tempting to define z as the energy savings times the price of energy (i.e., z would therefore represent the cost savings due to a change in energy use). However, a DSM program does not change the energy savings associated with the conservation measure--it changes the monetary cost of undertaking that conservation action. In other words, individuals who undertake the conservation action would receive the value of the energy savings whether or not the action was taken as part of a DSM program. However, the DSM program reduces the cost of undertaking the conservation action.

Consider a few examples. Suppose, as is the case with this paper, the program under consideration is an audit program. In this case, the cost reduction due to the program would be the cost of the audit. In other situations, the cost reduction is not so clear-cut. For example, suppose the program under consideration is an appliance rebate program, where the amount of the rebate varies with the efficiency of the appliance. In this case, the z variable is endogenous (i.e., the participant determines the actual rebate level he/she will receive). In such a situation, it is necessary to specify an additional equation for the program-induced cost savings.

The dependent variable (the y variable in Equation (2)) can take many forms depending on the characteristics of the program. For example, in an audit program the dependent variable is a binary variable which equals one if that household had an energy audit, and zero if not. In this situation, a logit or probit model is employed. In contrast, for a loan program the dependent variable can be the amount of the loan which was undertaken for conservation spending. Since this is a truncated continuous variable (i.e., there are no observations below zero), a Tobit model is used.

The dependent variable in this model considers everyone who took the conservation action, rather than just those who formally participated in the DSM program. According to Equation (1), these should be the same variables, but this may not be the case, for example, if some customers were unaware of the program. It is these non-participants (who took the action but are unaware of the program) who are key in the identification of this model. In practice, there is frequently enough of these individuals to provide the necessary information in most evaluations.

For those individuals who undertook the conservation action but did not formally participate because they were unaware of the DSM program, the cost savings of the program, z, did not affect their net utility. Therefore, the program effect is zero for these individuals, while it is equal to the cost reduction of the program for participants and non-participants who did not undertake the conservation action. A convenient method to incorporate this variation in the program effect is to define a dichotomous variable which is equal to one if that individual had heard of the program, and zero if not. In this design, the estimation of the behavioral model will interact this variable with the program effect variable.

Estimation

This section presents an example of how the above technique was used to determine the level of free riders in a residential home-energy audit program. The audit program in question was conducted by a Mid-western utility during the late 1980s. The data used in the analysis was obtained by mail surveys of participants and non-participants, with sample sizes of 307 and 381, respectively.

Determination of the level of free riders in an audit program is complicated because the main concern is not individuals who would have had an energy audit without the program, rather it is individuals who would have installed the conservation measures recommended by the auditor without the program. Therefore, a two-equation model is necessary, one to determine audit free riders and another to determine conservation-action free riders.

For the audit equation, the cost reduction of the program, z_1, is the normal cost of an audit (assumed to be \$50). Therefore, z_1 is equal to zero for non-participants who are not aware of the program, and z_1 equals the cost of an audit for participants and non-participants who are aware of the program.

For the action equation, the impact of the program, z_2, is incorporated by the inclusion of a dichotomous audit variable which is equal to one if the household had an audit and zero otherwise (irrespective of whether this audit was under the program in question or not). Since an audit provides information on how to save energy by undertaking specific conservation measures, one can view this variable as the decrease in informational costs associated with potential conservation actions.

Table 1 presents this estimated two-equation model. Because these two equations (audit and action) are interrelated, estimating each model separately will ignore the correlation between the two equations, and it would be inefficient. The proper approach to estimating this two-equation model is to use a simultaneous equation regression. Therefore, this model was estimated using the

Table 1. Bivariate Probit Estimation of Free Riders in an Audit Program

Variable	Audit Decision Coefficient (t-value)		Action Decision Coefficient (t-value)
Constant	-2.01 (-1.27)		2.821 (1.77)
Z_1	31.68 (5.67)		—
Z_2	—		0.586 (1.24)
Age of House	0.004 (1.05)		.005 (1.38)
Thermostat Setpoint at Night	0.029 (1.26)		-0.035 (-1.59)
Income ($000)	-.001 (-.238)		-0.006 (-1.12)
College Graduate	0.328 (1.27)		-0.089 (0.34)
Changed Square Footage	0.916 (3.621)		0.563 (1.84)
Will Own in Two More Years	-0.151 (-1.10)		-0.128 (-0.91)
Number of People Home at Night	0.114 (1.40)		0.017 (0.20)
Uses Gas for Space Heating	0.419 (1.99)		-0.038 (-0.16)
Number of People Aged 45 to 65	-0.103 (-0.78)		0.088 (0.67)
Number of People Under 5 Years Old	0.087 (0.38)		0.389 (1.46)
Replaced or Added a Major Appliance	0.191 (0.821)		0.212 (0.90)
Number of Years Owned Home	-0.004 (-0.43)		-0.015 (-1.51)
Correlation		-.131 (0.34)	

Bivariate Probit Model. This is a model of two simultaneous Probit equations, and it explicitly incorporates the correlation between the two equations. For a discussion of the Bivariate Probit Model, see Greene (1990).

The independent variables included in this model are:

- The age of the house;
- The thermostat setpoint at night;
- The household income (in thousands of dollars);
- A dichotomous variable denoting whether or not the head of the household is a college graduate;
- A dichotomous variable denoting whether or not the square footage of the house has changed since 1988;
- A dichotomous variable denoting whether or not the survey respondent plans to live in the house two years from now;

- The number of people home at night during the winter in 1990;
- A dichotomous variable denoting whether or not gas is used for space heating;
- The number of household members between ages 45 and 65;
- The number of household members under the age of 5;
- A dichotomous variable denoting whether or not a major appliance was replaced or added since 1988; and
- The total number of years the respondent has owned the house.

The choice of these independent variables was based on the a priori expectation of the variables that were thought to influence both the decision to undergo an audit and to take a conservation action.

The z_1 variable is highly significant and has the correct sign, and the z_2 variable has the correct sign but is not significant. Many of the variables in the model are not significant, but it was decided to leave them in to avoid omitted variable bias. Omitted variable bias occurs when a relevant independent variable is not included in the regression equation. In general, if this variable is correlated with the variables in the model, the estimates of the remaining variables in the model will be biased. This is contrasted with the inclusion of irrelevant variables, in which case the estimates are unbiased. This is particularly true of the correlation between the two equations, whereas the estimation results seem to indicate that the correlation is not significant. In addition, some of the variables do not have the expect sign. For example, the thermostat setpoint at night is inversely related to the probability of undertaking a recommended action. In most cases, these variables were not significant so the unexpected sign was not a major concern.

The derivation of the level of free riders in this audit program is found by using the estimated model to simulate whether or not an audit and conservation action would have been undertaken without the program. This amounted to determining the probability of an audit and action under the assumption that the household is unaware of the Audit program ($z_1 = 0$). This represents the free-rider level for both audit and conservation actions.

Table 2 presents the free-rider level for a typical participant. That is, the value for the independent variables are close to the mean value of participants. The resulting estimate of the level of audit free riders is 47%, and the level of action free riders is 52%. From the utilities' perspective, the important free-rider level is that for taking action, thus the 52% figure represents the level of free riders for this program. This figure compares favorably to the 58% of survey respondents who stated they would have undertaken the recommended action without the audit.

This procedure indicates another benefit of a behavioral model. Since the level of free riders depends upon the value of the independent variables, it becomes possible to conduct "what if?" scenarios by changing the values of these variables. For example, if the marketing changes focus by targeting high income household, the value for that variable in Table 2 can be changed to reflect the change in the program, resulting in a new estimate of the level of free riders given this change.

Conclusion

This paper presented a statistical behavior model of free ridership. This is a potentially useful approach to the problem of measuring free riders because it does not have the biases and subjective errors that can occur by using other methods.

A behavioral model also offers program planners the ability to simulate the level of free riders that can be expected to be associated with changes in participant and program characteristics.

However, because these behavioral models are a relatively new technique to determine the level of free ridership in a DSM program, many of the issues have yet to be completely resolved, and these models have seen only limited use and acceptance. In addition, these models require a large amount of information regarding non-participant actions. This information includes, for example, the efficiency of the purchased appliance (for a rebate program analysis). It is also possible that none of the non-participants undertook the conservation action in question, in which case it may be impossible to estimate the model.

Finally, the results of these models may be influenced by free drivers (i.e., non-participants who are affected by the program, even though they did not directly participate in the program).

Table 2. Estimation of Free Riders in an Audit Program

Variable	Mean Value for Participants
Age of House	50 years
Thermostat Setpoint	65 degrees
Income ($000)	$33,000
College Graduate	No (0)
Changed Square Footage	Yes (1)
Will Own in Two Years	Probably (2)
Number of People Home at Night	3
Uses Gas for Space Heating	Yes (1)
Number of People Aged between 45 and 65	1
Number of People Under age 5	1
Replaced or Added a Major Appliance	Yes (1)
Number of Years Owned Home	10
Audit Free Riders	47%
Action Free Riders	52%

References

Greene, W. H. 1990. *Econometric Analysis*. Macmillan, Inc., New York, N.Y.

Kreitler, V. 1990. "Market Issues in Free Rider Estimation." *Proceedings of the Electric Utility Marketing Research Council EUMRC/EPRI Conference*, pp. 1028. EPRI CU-7010, Bala Cynwyd, Pennsylvania.

Regional Economic Research, Inc. 1990. *SDG&E Non-Residential Audit Evaluation*. San Diego, California.

Saxonis, W. 1991. "Free Riders and Other Factors That Affect Net Program Impacts. *Handbook of Evaluation of Utility DSM Programs*, pp. 119-133. Oak Ridge National Laboratory, ORNL/CON-336, Oak Ridge, Tennessee.

Train, K. 1988. "Self-Selection in a New Context: Estimating the Impact of Conservation Programs on Measure Adoption." *Proceedings of the 1988 Summer Study on Energy Efficiency in Buildings*, Volume 9. American Council for an Energy-Efficient Economy, Washington, D.C.

Violette, D., and M. Ozog. 1989. "Correction for Self-Selection Bias: Theory and Application." *Energy Program Evaluation: Conservation and Resource Management Conference*, pp. 241-250. Argonne National Laboratory, Argonne, Illinois.