Who Dunnit? Determining Savings Attribution When Both Rebates and Financing Are Available

Stephen Grover, Evergreen Economics, Portland, OR
Jenny Fraser, Evergreen Economics, Portland, OR

ABSTRACT

California is experimenting with a range of different pilot programs that offer financing options for purchases of energy efficient upgrades and equipment. These financing programs operate in areas where rebates are also available for the same efficiency measures. From an impact evaluation standpoint, disentangling the influence of both rebates and financing on customer equipment choices presents a challenging impact analysis problem.

This paper presents a discussion of two different analysis methods that could be used to determine attribution of finance and rebate options in areas where both are available to customers. Analysis methods discussed include a self-report analysis approach and a discrete choice nested logit model.

The advantage of the self-report approach is that it is relatively easy to administer and can be used for most program situations. The disadvantage is that it is subject to potential biases due to issues of respondent recall and asking them to speculate on what they would have done in absence of the program. The nested logit model, on the other hand, is based on observed data that reflect customers' actual efficiency decisions in the market, thereby avoiding any biases relating to having survey respondents speculate on what they would have done in absence of the program. The primary disadvantage is that the discrete choice model has relatively high data requirements.

The paper concludes with recommendations on how best to structure the nested logit model and provides some important modeling considerations that need to be addressed prior to designing the data collection activities.

Introduction

A common approach for estimating net impacts of energy efficiency programs is to use a battery of survey questions to create a self-report estimate of free ridership. This method has been traditionally applied to rebate programs and is now being applied to energy efficiency financing as finance programs become more widespread. With the self-report method, program participants are asked a series of questions relating to their equipment purchase decision-making process to determine the influence that the program had in ultimately getting the customer to make an energy efficient equipment purchase. Questions typically focus on when the customers first became aware of the program, the influence the program had on the efficiency of the equipment chosen and the timing of the project. At some point the respondent is asked about what they would have likely purchased if the program assistance (e.g., rebate, financing, etc.) had not been available. A well-designed self-report survey will also include questions that allow for a consistency check of responses, to determine if respondents are providing logically consistent responses across related questions. These 'stated preference' survey data are then scored to develop a measure of free ridership and/or spillover and to calculate a net-to-gross ratio.

The self-report method has the advantage of being relatively easy to implement and can be applied to virtually any measure promoted as part of an energy efficiency program. A weakness of the self-report approach is the potential for biased results, as it relies on respondents remembering equipment purchase decisions that occurred in the past. Additionally, for commercial projects the purchase decision may have occurred over months or even years, which makes disentangling the influence of a utility rebate or finance program from other possible influences especially challenging.

One alternative to the self-report method is a type of discrete choice model known as a nested logit. The nested logit model can provide an estimate of the probability of a customer making a high efficiency equipment purchase relative to other reasonable choice alternatives. This model is based on 'revealed preference' data that reflect purchases that the customer actually made under observable market conditions. Because it is the one method that relies on revealed preference data that can be used to simulate purchases with and without rebate and financing programs in place, it offers an important alternative to the self-report method and therefore provides the focus for this paper.[1]

While the nested logit model has the advantage of relying on actual market data, it has substantial data collection requirements, and its application is limited to certain measures or types of projects that can be presented as discrete choices. Because of these limitations, the nested logit model is used much less frequently then the self-report method for estimating net impacts for energy efficiency programs.[2]

The remainder of this paper provides a relatively concise comparison of just two net-to-gross methods that were chosen to represent both stated preference and revealed preference data. The self-report method is the most commonly used approach and therefore provides a logical starting point for comparison. Other candidate approaches such as latent class discrete choice analysis and the analytic hierarchy process are less commonly used and have similar weaknesses as the other stated preferences methods. Consequently, they provide a less useful comparison. For analysis methods using revealed preference data, the nested logit model is a powerful statistical model that has not seen as much widespread use in energy efficiency program evaluations, which is likely due to the amount of data typically needed to estimate a model. Despite this challenge, we believe that the nested logit model should be more widely considered and is an especially promising method for disentangling the effects of program rebates and financing. For these reasons, we have chosen the nested logit as the second analytical method for comparison against the self-report method.

Framing the Research Question

Before choosing between the self-report approach and the nested logit model, it is critical that the overarching research objective be clearly identified. This is especially important for the nested logit model, where data collection can be expensive and the data needs will vary significantly based on the research question being addressed. For the purposes of this paper, we focus on two related (but different) attribution questions:

1. What share of energy efficiency program participation can be attributed to financing?
2. How does the availability of financing (both utility-sponsored and non-utility) affect the overall number of energy efficient purchases?

---

[1] Other variations on the discrete choice method that are being explored in finance program evaluations include latent class discrete choice and the analytic hierarchy process (AHP), but both of these approaches are currently less common than the self-report method. These methods also rely on self-report data, rather than the preferred revealed preference data used for the nested logit model.

[2] A nested logit model was used in the evaluation of California's Small Commercial Programs, see Itron (2010). For another example of a nested logit model applied to an impact evaluation, see Seiden and Platis (1999), where a nested logit model is used to estimate free riders and free drivers in a gas furnace program.

The first question is relatively limited in scope, as the research is directed toward allocating a set number of equipment purchases between different combinations of rebate and finance incentive offers. In other words, the question is asking what share of total purchases can be attributed to utility-offered financing versus rebates. This type of framework is consistent with a traditional impact evaluation, where the focus is on determining savings and attribution for an observed cohort of participants.

The second question has a broader perspective – it asks how the availability of financing is affecting the entire market for energy efficiency. In other words, does financing increase the total number of purchases in addition to increasing the number that can be attributed to an efficiency program? This can be further refined to determine the relative effect that utility-sponsored financing has on efficient purchases compared to the effect from financing obtained from non-utility sources.

How the self-report approach and nested logit model can be applied to each of these research frameworks is discussed below. Because the self-report approach is relatively common, the majority of the discussion focuses on the nested logit model and how it can be adapted for use in markets where both financing and program rebates are available for energy efficient purchases.

Self-Report Approach

The self-report approach is one of the most popular analysis methods used to estimate net impacts of efficiency programs. Through the self-report method, participants in an efficiency program are asked a series of questions regarding their energy efficient equipment purchases. These question batteries focus on the role of the program in determining what equipment was purchased and the decision to participate in the efficiency program. Ultimately, the self-report method attempts to understand what the customer might have purchased had the efficiency program (and the financial assistance provided by the program) not been available.

For a finance program, the self-report method can be tailored to ask detailed questions about the available financing options. This is an important advantage as it allows for questions that can determine the influence that financing had on the final decision to install energy efficient equipment. Note that this includes the initial decision on whether or not to install any efficient equipment, as well as the final design of the project. One of the key target areas for financing programs in California is whole-house retrofits, and the availability of financing for these projects may be leading to larger and more efficient projects than what would have been done otherwise. The self-report method provides an opportunity to measure just what influence the financing is having on the size and design of these projects. This flexibility is one of the key advantages of the self-report method, as questions can be designed to match virtually any type of program

The self-report method can also help determine program influence on the supply side or upstream part of these transactions. It may be that there are instances where the availability of financing is influencing the customer decision in ways in which the customer is unaware. The availability of financing and contractor training, for example, may lead contractors to develop easy financing packages for customers so that the financing application process is very simple from the customer standpoint. The customer may not remember financing as being particularly influential if the application process only lasts a few minutes. The contractor, however, may view the availability of financing as a critical component in closing the deal. Administering a self-report survey for contractors provides a way to capture this important program interaction and incorporate it into the net impact analysis.

Another important advantage of the self-report method is that it is a relatively inexpensive method for determining net program impacts. In cases where there are no market data showing the alternatives to the program, the self-report method may be the only viable option. The primary drawback of this method is the reliance on stated preference data. Because it is asking the respondent to speculate on what they would have done had the program not existed, there is no way to verify the accuracy of these responses,

With the self-report method, one topic of considerable debate has been designing the appropriate questions to minimize potential biases. Types of biases include recall bias, where the respondent is not able to remember accurately and consequently provides inaccurate information on their purchases to the interviewer. Social desirability bias is also a concern, where respondents answer questions in a way that makes them appear in a more positive light. For example, respondents may claim that they would not have made their purchase without the rebate, when in actuality they had already made up their mind on what they were going to purchase and accepted the rebate anyway, even though it had no influence on their decision.

An additional issue that has received considerable attention is determining how responses should be weighted to develop a single measure of free ridership. Since historically there has been no consensus on consistent methods, comparisons across programs (and even evaluations of the same program over different evaluations) have been difficult.

To help impose some consistency on how the self-report method is applied, the California Public Utilities Commission (CPUC) Energy Division convened a committee of evaluators to develop a standard framework for the measurement of net-to-gross ratios for residential and small commercial programs in a systematic and consistent manner using the self-report approach. With the assistance of its technical consultants and evaluators, the Energy Division developed the *Guidelines for Estimating Net-To-Gross Ratios Using the Self-Report Approaches* (Guidelines), which provides more detailed guidance than was available in the earlier California Evaluation Protocols.

Programs that provide both financing and rebates present additional challenges for all impact estimation methods, including the self-report approach. The incremental effect of financing when a rebate is also applied is difficult to determine when both a rebate and financing are used for the same equipment purchase. In some cases, these two types of incentives may be discussed together with the installation contractor, making it virtually impossible to determine the influence of either component individually. The issue is further compounded by the fact that financing is available both through the efficiency program and through outside sources. Self-report method questions then must distinguish between sources of financing in order to develop a clear idea of the importance of financing that is available *only through the program.* Finally, given the importance of contractors, an approach that incorporates the contractor influence should be considered in order to obtain a complete picture of the role that financing has on the customer's ultimate purchase decision. This may necessitate a separate battery of self-report questions that are directed to contractors to determine the importance that financing plays in both getting customers to choose energy efficient equipment and expanding projects that eventually result in greater energy savings.

In summary, the self-report approach's reliance on stated preference data is both a blessing and a curse. While the method allows for an efficient means of estimating net impacts, concerns about biased data and arbitrary weighting schemes make the self-report method results less reliable. An alternative that helps avoid these issues is the nested logit model. While this approach is not without its own problems, it does have the advantage of relying on actual customer choices and therefore is less speculative in nature compared to the self-report approach.

Nested Logit Modeling Approach

The nested logit model combines customer information about their equipment choices and purchase decision process with information about measure costs and savings impacts to estimate the probability that an individual equipment option is purchased. Note that due to space considerations, a detailed statistical discussion of the nested logit model is not presented here, but can be found in standard statistics textbooks that address discrete choice modeling.[3] The discussion below instead focuses on the general framework for the nested logit model, followed by some important practical considerations that will help with setting up the data collection needed to estimate the model.

To simulate the choice process, the nested logit model uses a decision tree structure to represent the different decision stages. A decision tree that could illustrate possible choices with and without financing is shown in Figure 1. In this decision framework, the first decision level is whether or not any equipment purchase should be made. In the second decision stage, multiple equipment options are shown with and without financing and rebates. Each option designates a distinct purchase choice and would be defined in the model as a series of equipment or project characteristics that might include project cost, rebate, energy savings and financing (defined as interest rates, monthly loan payment, on-bill or off-bill, and/or other loan characteristics of interest).

Note that these choices could be expanded to split the finance options into separate standard and high efficiency equipment choices, but for ease of presentation, they are combined in the diagram. The finance options (Options A and B) could be further split to show whether financing is obtained through a utility-sponsored program or from non-utility sources. The choices could also be subdivided into standard and high efficiency options, which may allow for the possibility of estimating spillover. Option D still involves an equipment purchase, but the purchase is done without the assistance of a rebate or financing.

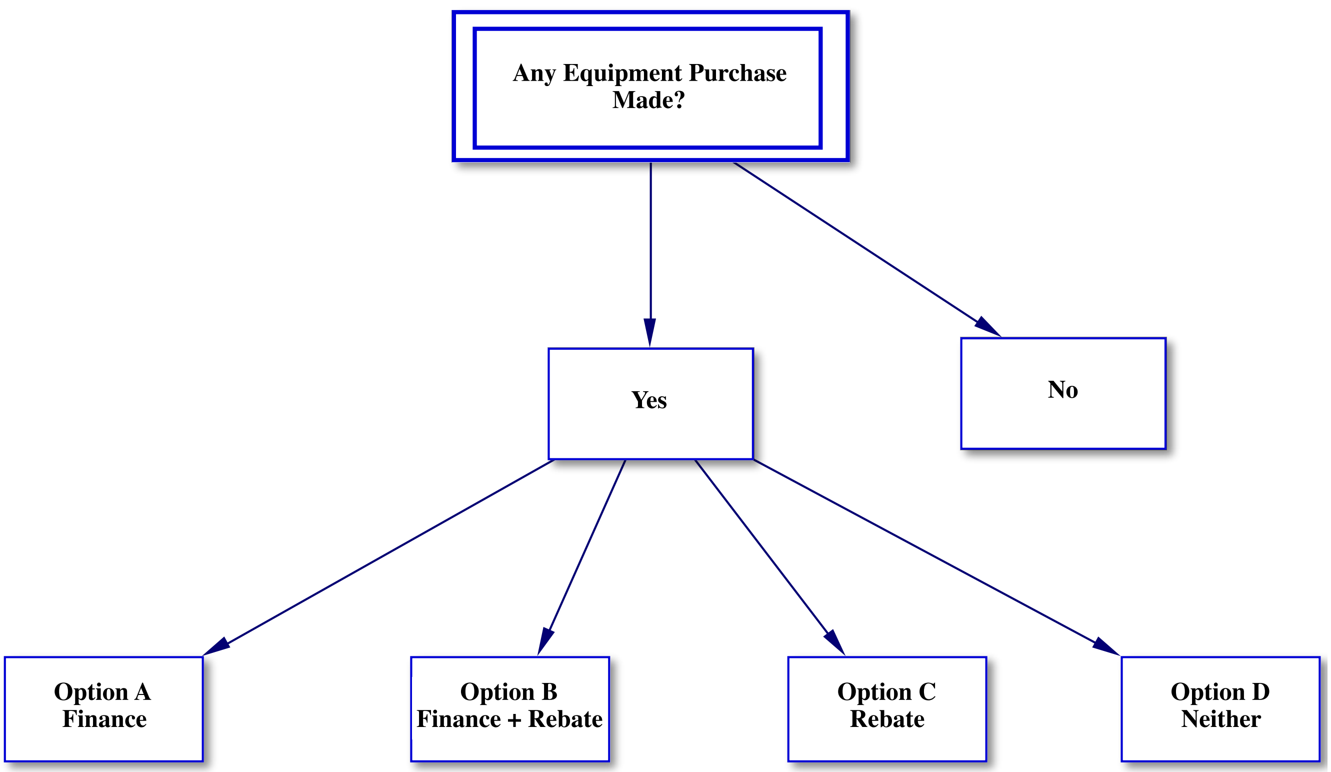

Figure 1. Sample Decision Tree for Nested Logit Model

[3] For a full technical discussion of the nested logit model specification, see *Limited-Dependent and Qualitative Variables in Econometrics* by G. S. Maddala, Cambridge University Press (1992) and *Econometric Analysis* by W. Greene (2011). The software program LimDep is designed specifically to estimate discrete choice models such as the nested logit, and the software manual also provides a good background on the statistical theory underlying the model.

As shown in this decision tree example, the nested logit model framework can be designed to include a no purchase option, which is consistent with Research Question #2 discussed at the beginning of this paper. With this framework, the model allows for entry into the market, as the model is estimating the effect that the program has on encouraging additional equipment purchases – not simply the choice between efficient and inefficient purchases for those who have already decided to purchase.

An important feature of the nested logit model is that it provides a way for the benefits of the lower stages of the tree to influence the earlier decisions. In this example, the attributes of the equipment[4] options in the second stage (which may include energy savings, rebates, and/or financing options) affect the decision of whether or not to make any equipment purchase at all (the first stage). Each decision stage is estimated with the relative benefits of each stage linked to the other stages through an "inclusive value" variable that represents the relative benefits of the lower model stages.[5] This term allows the perceived benefits from the lower parts of the decision tree to have an influence on the decisions made in the upper parts of the decision tree.

An additional advantage of the nested logit model is that it addresses the entire equipment choice decision for both program participants and non-participants using a structure that is consistent with standard microeconomic theory and utility maximization. Additionally, if the model includes both purchasers and non-purchasers, it eliminates the potential problem of self-selection bias that may occur if only data on purchasers or participants were used. Because the nested logit model simulates the entire decision to purchase energy efficient equipment explicitly and includes observations for customers who choose options outside the program, the problem of self-selection bias is avoided.

Given this analysis framework, the nested logit model can be used to estimate attribution provided that a decision tree model structure is designed that allows for the delineation of the rebate and financing effects. At a minimum, this requires data on customers making equipment purchases using rebates and financing (together or separately), and at least some of the purchases need to utilize no more than one of these options (e.g., some purchases use rebates or financing, but not both). If the rebates and financing are always used together, then they will be perfectly collinear and the model will not be able to disentangle the individual effects and determine attribution.

To estimate the model shown in Figure 1, additional data need to be collected on customers that have not made an equipment purchase. To be consistent with the model assumption of perfect information, a more robust model would collect information on non-purchasers that were aware of the utility rebate program and the financing options that were available to them both inside and outside the program. As this suggests, adding this additional non-purchasing cohort to the data collection will increase the data collection needs of the model substantially.

For a given decision tree structure, the probability of purchasing any given equipment Option A (high efficiency equipment with financing, in this example) can be expressed as the product of two separate probabilities: the probability that a purchase is made (either standard or high efficiency) multiplied by the probability that equipment Option A is chosen given that a purchase is made. These two probabilities relate to the two stages of the decision tree shown in Figure 1 and are estimated simultaneously in the nested logit model. In equation form, this can be written as:

$$\text{Prob(EE Purchase \& Equipment A)} = \text{Prob(EE Purchase)} \times \text{Prob(Equipment A} \mid \text{EE Purchase)}$$

---

[4] This paper refers to purchases as an 'equipment' choice, but the discussion could easily be expanded to include larger retrofit projects that include multiple measures or retrofits, as long as the whole project can be considered as a distinct discrete choice.

[5] See Maddala (1992) and Greene (2011) for a more technical discussion of the inclusive value variable.

To determine these probabilities, the nested logit model combines customers' choices of equipment with information on measure costs and savings impacts to estimate the probability that alternative equipment options will be chosen.

A general form of the equation that reflects the potential benefit of the equipment choice is as follows:

$$Equip\ Choice = \beta' Rebate + \beta' Finance + \beta' MeasureCost + \beta' Z + \varepsilon$$

The coefficients on *Rebate, Finance* and *MeasureCost* are used to define the characteristics of each equipment choice option. The remaining explanatory variables *Z* contain additional variables that reflect characteristics of the equipment choices or characteristics of the customer. The dependent variable *Equip Choice* will have a value of either zero or one, where customers are given a value of one to indicate their actual equipment choice or a zero for all non-chosen alternatives.

Once the model is estimated, the coefficient estimates are combined with data on the various equipment choice options to determine the probability that any of the purchase/equipment options are chosen. These probabilities are then used to simulate the change in customer behavior based on changes to program-related variables such as financing and rebates.

After the probabilities are calculated using the nested logit probability formulas, the net-to-gross ratio can be calculated that estimates the portion of energy efficient installations that can be attributed to utility-sponsored financing. This is done by calculating the change in probability of purchasing high efficiency equipment with and without the financing options available. In these calculations, the 'no program' scenarios are simulated by setting any program-related variables (e.g., rebate amount, utility-sponsored financing, program awareness, etc.) equal to zero and then recalculating the choice probabilities. As shown in the formula below, the net-to-gross ratio is the difference in the probability of purchasing energy efficient (EE) equipment with and without utility-sponsored financing divided by the probability of purchasing the energy efficient option when the utility-sponsored financing is available:

$$NTG = \frac{\text{Prob}_{Totalj}^{W/PROG\_FIN} - \text{Prob}_{Totalj}^{WO/PROG\_FIN}}{\text{Prob}_{Totalj}^{W/PROG\_FIN}}$$

Where:

$\text{Prob}_{Totalj}^{W/PROG\_FIN}$ = Probability of choosing EE equip option j WITH the Financing Program

$\text{Prob}_{Totalj}^{WO/PROG\_FIN}$ = Probability of choosing EE equip option j WITHOUT the Financing Program

Depending on how the model and net-to-gross ratio are defined, the calculation can be modified to address different questions regarding the role of financing. For example, the analysis can be designed to address Research Question #1 discussed at the beginning of this paper, where one wants to estimate the role that utility-sponsored financing has on program participation (e.g., program participation attribution needs to be allocated between rebates and financing). In this case, the results of the nested logit model could be used to estimate the net-to-gross ratio as follows:

$$NTG = \frac{\text{Prob}_{Totalj}^{W/PROG\_FIN} - \text{Prob}_{Totalj}^{WO/PROG\_FIN}}{\text{Prob}_{Totalj}^{W/REBATES\&PROG\_FIN} - \text{Prob}_{Totalj}^{WO/REBATES\&PROG\_FIN}}$$

Where :

$\text{Prob}_{Totalj}^{W/PROG\_FIN}$ = Probability of choosing EE equip option j WITH the Financing Program

$\text{Prob}_{Totalj}^{WO/PROG\_FIN}$ = Probability of choosing EE equip option j WITHOUT the Financing Program

$\text{Prob}_{Totalj}^{W/REBATES\&PROG\_FIN}$ = Probability of choosing EE equip option j WITH the Financing and Rebate Program

$\text{Prob}_{Totalj}^{WO/REBATES\&PROG\_FIN}$ = Probability of choosing EE equip option j WITHOUT the Financing and Rebate Program

In the above equation, the denominator is the total change in probability of purchasing high efficiency equipment due to both the utility rebates and utility-sponsored financing program. The numerator is the share of the change that can be attributed to just the availability of the utility-sponsored financing.

A modified version of this calculation can be used to determine the role that just the utility-sponsored financing is having on high efficiency equipment purchases. If the decision tree and model are set up with separate choices for utility-sponsored and non-utility financing, then the influence of the utility-sponsored financing can be estimated by setting just the program finance variables equal to zero:

$$NTG = \frac{\text{Prob}_{Totalj}^{W/REBATES\&PROG\_FIN} - \text{Prob}_{Totalj}^{WO/PROG\_FIN}}{\text{Prob}_{Totalj}^{W/REBATES\&ALL\_FIN}}$$

Where :

$\text{Prob}_{Totalj}^{W/REBATES\&ALL\_FIN}$ = Probability of choosing EE equip option j WITH the Rebate Program and all finance options (utility-sponsored and non-utility)

$\text{Prob}_{Totalj}^{WO/PROG\_FIN}$ = Probability of choosing EE equip option j WITH the Rebate Program but WITHOUT utility-sponsored financing

Finally, the above equation can be expanded to address Research Question #2 and estimate the effect that all financing options (both utility-sponsored and non-utility) have on the likelihood of purchasing high efficiency equipment:

$$NTG = \frac{\text{Prob}_{Totalj}^{W/REBATES\&ALL\_FIN} - \text{Prob}_{Totalj}^{WO/REBATES\&NO\_FIN}}{\text{Prob}_{Totalj}^{W/REBATES\&ALL\_FIN}}$$

Where :

$\text{Prob}_{Totalj}^{W/REBATES\&ALL\_FIN}$ = Probability of choosing EE equip option j WITH the Rebate Program and all finance options (utility-sponsored and non-utility)

$\text{Prob}_{Totalj}^{WO/REBATES\&NO\_FIN}$ = Probability of choosing EE equip option j WITH the Rebate Program but WITHOUT any financing opportunities (utility-sponsored and non-utility removed)

As the above examples demonstrate, with careful design the nested logit model can be used to address a wide range of net impact and attribution questions, including how the availability of financing (both inside and outside of utility programs) can influence both the number of energy efficient purchases and the overall size of the energy efficient market in terms of energy savings. The nested logit model also has the advantage of relying on actual customer purchases rather than stated preference data. The primary drawback of the nested logit model is the amount of data required, particularly for models that incorporate the decisions of non-participants. Due to the difficulty of collecting purchase activity for customers outside utility programs, the effort required to survey these customers might be cost prohibitive.

Additional considerations for designing and implementing a nested logit model are provided below.

Guidelines for Specifying the Nested Logit Model

There is a wide range of possible decision tree structures available for the nested logit model, and the specific structure will be determined both by the research question being addressed and the availability of the data needed to estimate the model. With any of these structures, however, there are common factors that need to be considered before finalizing the model design and beginning the data collection needed to support the model. These considerations will influence the amount and type of data that need to be collected, which in turn may affect what type of model specification is ultimately chosen.

Consideration #1: Variables that do not vary across choices will drop out of the model. In the logit formula used for this model, variables that do not change in value across choices will cancel out in the final model. Possible instances where this might occur are with any customer-specific variables (e.g., awareness, household size, home vintage) that will not vary across equipment choices. If equipment characteristics do not vary across choices (e.g., interest rates, rebate amounts), then these will also drop out of the model. This issue can be resolved through careful construction of the variables that allows for variation across choices.

A related issue is collinearity across choice variables, which can be an issue if the value of one variable is a linear function of another. For example, it could be that a rebate is calculated as a set percentage of total project cost, in which case project cost and rebate would be collinear. A similar situation could arise if the financing variable is expressed as finance cost, which is a function of total project cost. Careful construction of choice-specific variables can get around many of these problems. Having project cost and rebate combined into a single net project cost variable is one such example.

Consideration #2: Values for all choices must be provided or imputed for each customer decision opportunity. The underlying economic model assumes that each customer has perfect information about their equipment choices and then makes an informed choice after comparing the relative costs and benefits of each. For this reason, the model must impute all the equipment option characteristics (e.g., price, savings, rebate, financing terms, etc.) for all alternatives even if none were chosen. This can be done using average values for different customer classes that actually selected each option, where the market data are known.

Note that this applies to models that include a no purchase option (Figure 1), as the consumer choice theory assumes that the decision not to purchase is made after considering the other purchase options available. In reality, most non-purchasers are unlikely to be actively considering these options at any particular moment, which presents a challenge for data collection, as non-purchasers may not be able to provide accurate information on what they would likely purchase. In order to make the model more consistent with the underlying assumptions, we recommend that data collected for non-purchasers be collected from customers that have been considering a purchase within the last year but have not yet made a purchase.

Consideration #3: Customer awareness of program options is an important variable that should be included in the model. Simple awareness of a program can be a key driver of participation, and therefore should be included in the model. It also helps incorporate into the model the effects of contractor marketing, which is a common source of program awareness among customers; consequently, many programs rely heavily on contractors to market the program directly to their customers.

Program awareness can be incorporated into a nested logit model in several ways. It can be included as a separate stage in the first stage of the model (depending on how the decision tree is structured). This is accomplished by interacting the awareness variable with choice-specific constants, so that there is variation across options and the variables do not cancel out in the logit equation. It can also be incorporated in the second stage through the values of other program-related variables such as rebates and financing. If the customer is unaware of the program, then the values for rebates and financing would be set to zero for all program options for that customer.

Consideration #4: The coefficient estimate for the inclusive value term provides information on if the decision tree structure is correctly specified. An important aspect of the nested logit model is that options in the lower levels of the decision tree are grouped into similar categories. This is done to avoid irrelevant options having an inappropriately large effect on the other choice probabilities (e.g., the Independence of Irrelevant Alternatives (IIA) property). Clustering similar options into groups in the lower levels of the tree avoids this problem if the grouping is done properly. For example, choices may be grouped by efficiency level, end use or financing source as long as the groupings can reasonably be considered to contain closer substitutes within groups than across groups.

One of the advantages of the nested logit model is that the coefficient on the inclusive value term that links different stages of the model provides an indication of whether or not the chosen grouping is appropriate (e.g., choices are closer substitutes within groups than across groups). With a correctly specified decision structure, the inclusive value coefficient will have a value between zero and one. If the coefficient lies outside this range, then an alternative grouping of choices should be explored.

Consideration #5: Neither the nested logit model nor the self-report approach is likely to provide reliable results if customers are always using both financing and rebates for equipment purchases. To understand how choices vary with rebates and financing, the analysis dataset needs to have at least some observations where customers only use financing or only use rebates. Note that having observations where both financing and rebates are used is acceptable, provided that there are some observations that use only one of these options so that enough variation is included in the model and all of the nodes of the decision tree are represented in the data. If the analysis is being done on a program that requires customers who want financing to also use a rebate, then (at a minimum) the dataset must also include some customers who chose the rebate only option (e.g., made a high efficiency purchase using a rebate but without financing).

Without some variation in the use of rebates and financing across customers, neither method is likely to disentangle the relative influence of either incentive. In the nested logit model, rebates and financing will be so highly correlated that the model will not be able to allocate the influence to the two variables individually. Similarly, customers will be considering the rebates and financing simultaneously as they make their equipment choice and consequently, even a carefully constructed questionnaire is unlikely to be able to distinguish the incremental effect between these two factors. In order for either method to work, the analysis dataset should include some customers who only used rebates and other customers who only used financing.

Consideration #6: Most model specifications will need to have data on non-participant equipment purchases. The type of non-participant equipment purchases will depend on the research question being addressed, but at a minimum, the model will usually require information on standard efficiency purchases done outside a utility efficiency program. To address the importance of financing, non-participant purchases should also include some high efficiency options that have been financed using non-utility sources (home equity loans, credit cards, etc.). Having these data provides a fuller picture of how well program factors such as financing, rebates and program awareness are influencing the number of purchases as well as efficiency levels. The data collection needed to identify these types of purchases in the general population where tracking data are not available will add to the cost of a project, as a significant survey effort is needed to first identify the non-participant purchases and then recruit them to complete the survey.

References

Greene, W. H. 1990. *Econometric Analysis*. New York: Macmillan Publishing Company.

Grover, S., T. Helvoigt, and J. Yaillen. 2011. "Free to Choose? A Comparison of a Nested Logit Model with a Billing Regression Model and Self-Report Analysis in a Commercial Impact Evaluation." *In Proceedings of the 2011 International Energy Program Evaluation Conference.* Boston, Mass.: International Energy Program Evaluation Conference.

Itron. 2010. *Small Commercial Contract Group Direct Impact Evaluation Report (CALMAC ID CPU0019.01).* Prepared for the California Public Utilities Commission.

Maddala, G. S. 1992. *Limited-Dependent and Qualitative Variables In Econometrics.* New York: Cambridge University Press.

Ridge, R., K. Keating, L. Megdal, and N. Hall. 2007. *Guidelines for Estimating Net-To-Gross Ratios Using the Self Report Approach.* Prepared for the California Public Utilities Commission.

Seiden, K. and H. Platis. 1999. "Freerider and Freedriver Effects from a High-Efficiency Gas Furnace Program." *In Proceedings of the 1999 International Energy Program Evaluation Conference.* Denver, Colo.: International Energy Program Evaluation Conference.

TecMarket Works. 2006. *California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals.* Prepared for the California Public Utilities Commission.