Market Penetration: How to Predict the Future

Stephen Grover, Ibbie Ikenze-Bates
Quantum Consulting Inc., Berkeley, Calif.

ABSTRACT

One of the biggest challenges in evaluating energy efficiency programs is predicting how customers will react to future changes in incentives and technology. This is especially true within a competitive energy market. This paper presents a market penetration model based on stated preference experiments that is designed to address this issue. This model can be used to predict customer purchases under alternative market conditions such as changes in technology, program rebates, and program qualifying equipment.

The stated preference experiments elicit ratings from residential customers on program and equipment attributes such as price, rebate, and annual energy savings. The data were collected and the model estimated using information from Florida Power and Light's Residential HVAC Program. Once the data are gathered, a logit model is estimated to determine the probability that each program and equipment option is chosen. The model is calibrated to actual customer purchases and then used to predict future equipment purchases and program participation. When values for very high efficiency equipment are included in the experiment, the model can be used to forecast future purchases even when purchases of these units are not currently widespread.

This market penetration model provides a method to forecast equipment purchases, while taking into account future changes in technology. This model is flexible enough to incorporate new program features and changes in technology, and, as a result, will be valuable to any utility seeking to continue energy efficiency programs in a competitive market.

Introduction

One of the biggest challenges in evaluating energy efficiency programs is predicting how customers will react to future changes in technology and program incentives. Changes in technology, customer preferences, and demographics all influence the effect of energy efficiency programs. The challenge of planning for the future is compounded within a competitive energy market, where future market conditions are especially uncertain and new programs and products from competing firms are being introduced.

Conjoint analysis is a method that is especially well suited to address market penetration estimation challenges. Conjoint analysis involves conducting controlled stated preference experiments where people are presented with different product characteristics and are then asked to rank bundles of characteristics in order of their preference. This allows for hypothetical products or program scenarios to be evaluated relative to current market conditions. The information gained can then be used to predict how customers will react when the hypothetical products are introduced into the market.

Conjoint analysis was used in this manner to develop market penetration models for Florida Power and Light's (FPL's) Residential Energy Efficiency Programs. Models were developed for the following FPL programs: the residential HVAC Program, the Load Control or "On Call" Program, and the Duct Testing and Repair Program. Only the HVAC Market Penetration model will be presented in this paper as the analysis methods for each of these models are similar.

The purpose of the HVAC Market Penetration Model (the HVAC Model) is to provide a tool that can be used to predict future HVAC purchases both in and outside of FPL's HVAC Program. The final market penetration model is a self-contained computer software that con be used by FPL program managers to evaluate the current program as well as alternative program configurations and market scenarios. As a result, the HVAC Model is a powerful tool for developing a cost effective program configuration. The development of the HVAC Model as well as how it can be used to develop a more cost effective program is presented in the remainder of this paper.

The FPL Residential HVAC Program

FPL's Residential HVAC Program pays residential customers a rebate for installing high efficiency cooling equipment, either Central Air Conditioners (CACs) or heat pumps. To qualify for the HVAC program under the 1997 program standards, the installed cooling equipment must be have an efficiency rating of at least 11 SEER. Customers are paid a rebate based on both SEER rating and capacity. In addition, customers must not have participated in the program previously and must be on the FPL system for at least a year. To date, the Residential HVAC program has had more than 300,000 participants.

FPL is in the process of refining its energy efficiency programs and revising its program goals. Part of this process has involved developing market penetration modeling software that will predict future participation both under the current program standards and under possible alternative program configurations. The HVAC Model will be used by FPL to tailor the HVAC Program to meet the new program goals in the most cost-effective manner.

Conjoint Analysis

A stated preference technique known as conjoint analysis was used to collect data on customer preferences for use in the HVAC Model. Conjoint analysis is a well-established market research technique that is often employed to determine preferences for different products. It is an especially useful tool for evaluating products and services that have not yet been introduced into the market.

Conjoint analysis involves having respondents sort through and rank cards showing different product characteristics. Each card represents a different product choice and respondents are asked to rank cards in order of their preference. Since each card contains several characteristics of the product, respondents are forced to decide which characteristics are most important, and to make tradeoffs between different levels of product attributes when ranking the cards. For example, increased electricity savings comes at the cost of a more expensive air conditioning unit. Examining how these tradeoffs are made in the controlled conjoint experiment provides useful information on how customers evaluate these tradeoffs in the market when selecting their HVAC equipment.

Conjoint analysis has the advantage over revealed preference data of introducing hypothetical features into the analysis. For example, characteristics that currently do not exist in the market, such as different rebate levels or noncash incentives, can be introduced as equipment or program characteristics. Examining how these new features are evaluated in the conjoint analysis relative to the other equipment characteristics provides useful information on how these features will be evaluated in the market place.

For the stated preference experiments, customers were recruited from three different cities in FPL's service territory: Miami, Daytona, and Sarasota. As much as possible, the sample was divided between nonparticipants and those people who had already participated in FPL's HVAC program. The sample size from each city as well as equipment type is given in Exhibit 1. Customers were given either a CAC or heat pump deck to sort, depending on what type of cooling equipment they had in their homes. The CAC experiment involved ranking 16 different cards, the heat pump deck contained 25 cards for ranking. Respondents were asked to rank these cards from most preferred to least preferred as if they had already decided they were going to purchase a new cooling system and were now sorting through the available equipment options.

Exhibit 1
Sample Size by City and HVAC Program Participation

City	Participant Status	CAC	Heat Pump
Miami	Participant	83	0
	Nonparticipant	34	0
Daytona	Participant	53	76
	Nonparticipant	16	25
Sarasota	Participant	55	58
	Nonparticipant	23	15
Total		264	174

The attributes on the cards and the levels for these attributes are shown in Exhibit 2. Each card shows the price, savings, rebate, and financing for a particular CAC or heat pump. "Price" is the cost of the HVAC equipment without a rebate; "rebate" is the rebate offered by FPL off the purchase price of the equipment; "savings" is the average annual dollar amount of electricity savings for the customer; and "financing" describes different financing options available to the customer. The ranges given for these attributes reflect 3 ton HVAC units in the 10 SEER to 15 SEER range. Ranges were designed to cover possible future scenarios for the program, such as eliminating rebates for lower SEER options.

The attribute levels were randomly assigned across cards so that each of the attributes were perfectly uncorrelated within the deck. This is known as an orthogonal design, and has several advantages. By randomly assigning attribute levels, respondents are forced to make tough tradeoffs by deciding which attributes are most important. For example, in the market high electricity savings comes at the expense of higher equipment costs. If this were reflected in the cards with savings and price correlated, the difference in ranking due to either equipment price or savings cannot be determined from the data. Having the attribute levels randomly assigned in an orthogonal design mitigates this problem.

Another advantage of the orthogonal design is that it allows for preference functions to be estimated individually without the loss of precision that might occur when attribute levels are correlated. Estimating individual rather than aggregate utility functions using conjoint data has been shown to provide more accurate predictive results.[1] Since this analysis focuses on developing a forecasting model, individual utility functions were estimated and used as inputs for the HVAC Model.

Exhibit 2
Attributes and Attribute Levels for CAC and Heat Pumps Used in Conjoint Analysis

Card Attribute	Attribute Levels	
	CAC Trade	Heat Pump Trade
Price of Equipment (without rebate)	•$2000 •$2600 •$3200 •$4500	•$2200 •$3000 •$3800 •$4800 •$6000
FPL Rebate	•None •$200 off purchase price •$400 off purchase price •$800 off purchase price	•None •$200 off purchase price •$400 off purchase price •$800 off purchase price
Annual Electricity Bill Savings	•$50 •$200 •$400 •$550	•$50 •$200 •$400 •$600
Source of Financing	•Own •FPL •Manufacturer	•Own •FPL •Manufacturer

Statistical Model

The decision to purchase new HVAC equipment can be broken into two components; the probability of purchasing any HVAC equipment multiplied by the probability of choosing a particular HVAC equipment option given that the decision to purchase has already been made. In equation form

$$\text{Prob(Purchase \& J)} = \text{Prob(Purchase)} * \text{Prob(J | Purchase)} \quad (1)$$

Where Prob(Purchase & J) = The probability of purchasing and choosing equipment option J
Prob(Purchase) = The probability of making an HVAC purchase
Prob(J | Purchase) = The probability of purchasing equipment option J given that the decision to purchase has already been made.

[1] The issue of individual versus aggregate regressions in conjoint analysis is the topic of some debate. See, for example, Louviere and Hensher (1983) and Moore (1980) for a comparison of individual and aggregate estimation results using conjoint data.

The first half of equation (1) is the probability that an HVAC purchase is made. Based on past market research in FPL's service territory, the age of the old cooling system was found to be the only significant determinant in the decision to purchase equipment. Therefore, the HVAC model focuses only on replacements rather than all HVAC purchases. As a result, the HVAC Model uses the age of the existing system as the sole determinant of the likelihood of making an HVAC equipment purchase.

The second half of equation (1) is the probability that a particular equipment option J is chosen, given that the decision to purchase equipment has already been made. A random utility structure is useful for evaluating this equipment choice decision. Using the random utility model, the utility or benefit of any equipment option j for individual i is given as

$$U_{ij} = \beta' X_{ij} + \varepsilon_{ij} \tag{2}$$

Where U_{ij} = Total utility associated with choice j for person I
β = Utility coefficients or "part worths" to be estimated
X_{ij} = Vector of attributes for choice j for person I
ε_{ij} = Random utility component.

With the random utility model, utility is divided into an observed component ($\beta' X_{ij}$) and a random component (ε_{ij}) that captures the unobserved influences on the equipment choice decision. The distribution of the random component in (2) determines how the model is estimated. If the random component is distributed logistically, then the model can be estimated as a conditional logit.

With a random utility model, the deterministic portion of utility also determines the probability of any individual equipment option being chosen from the choice set.[2] Using the logit density function, the probability that person i chooses option j among n choices is

$$\text{Prob}(J \mid \text{Purchase}) = \exp(\beta' X_{ij}) \,/\, \Sigma_N \exp(\beta' X_{ij}). \tag{3}$$

This equipment choice probability is the second half of equation (2). Using equation (3) and the attribute levels from the conjoint cards, the probability of choosing each equipment option is calculated. The equipment choice set is constructed using the prices, savings, rebate, and financing associated with purchasing HVAC equipment in the 10 SEER to 15 SEER range as listed on the conjoint cards. Each customer has a choice set of 11 options with 6 equipment choices outside the program (10, 11, 12, 13, 14, 15 SEER) and 5 units within the program (11, 12, 13, 14, 15 SEER). The probability of choosing each equipment option is calculated at the individual level and then averaged across all of the individuals in the sample to get an overall probability of choosing each option.

One way to estimate the equipment choice probabilities is to estimate how each of the attribute levels on the conjoints cards affect the card ranking. The resulting coefficient estimates can then be used to estimate the equipment choice probabilities from equation (3) that are consistent with the random utility model in equation (4). Given the card attributes shown in Exhibit 2, the model to be estimated is

$$\text{Card Rank}_{ik} = \beta_{1i}'\text{PRICE}_k + \beta_{2i}'\text{SAVINGS}_k + \beta_{3i}'\text{REBATE}_k + \beta_{4i}'\text{FPLFIN}_k + \beta_{5i}'\text{MFGFIN}_k + \varepsilon_{ik} \tag{4}$$

[2] This link between the random utility model and the choice probability using the logit model was developed by McFadden (1973). See also Maddalla (1983), particularly pp. 60-61, for a concise derivation.

Where β_i = Parameters to be estimated for individual i

$PRICE_k$ = Equipment price shown on card k

$SAVINGS_k$ = Annual expected bill savings on card k

$REBATE_k$ = Rebate on card k

$FPLFIN_k$ = Dummy variable indicating FPL Financing on card k

$MFGFIN_k$ = Dummy variable indicating manufacturer financing on card k

i = Index for individual i

k = Index for card k

ε_{ik} = Random error term assumed to be logistically distributed.

This form of the logit model, where the dependent variable is the ranking of the card rather than the traditional zero or one value, is referred to as an exploded logit model. The exploded logit has the advantage over the conditional logit in that it utilizes all of the available ranking information.

Although the exploded logit model is slightly different than the traditional conditional logit model, the density function and the log likelihood function for the two models are identical. A behavioral interpretation of the exploded logit model using ranked conjoint data is offered by Allison and Christakis (1994). With a dependent variable as a ranking, the probability of each ranking can be interpreted as the probability that each successive ranked bundle is chosen given that the higher ranked bundles have already been selected and therefore removed from the choice set. The product of these probabilities yields the log likelihood function for the traditional conditional logit function and is consistent with the random utility specification. As a result, the estimation resulting from equation (4) can be used to estimate the equipment choice probabilities using equation (3).

Equation (4) is estimated separately for the CAC and heat pump trades. Since individual utility functions were estimated, it is not possible to display all of the estimation results. However, the average coefficient estimates across all respondents are given in Exhibit 3. In general, the estimations results are consistent with expectations. Equipment cost has a negative estimate while savings and rebate are positive for both CAC's and heat pumps. The heat pump estimate for savings is an order of magnitude higher than the heat pump estimate for rebate. This may be reflecting a greater desire for electricity bill savings amongst heat pump owners, who have shown that they are willing to pay to the higher equipment price for heat pumps to realize this savings. For the financing options, the coefficient estimates reflect the desirability of the financing option relative to the customers relying on their own financing. For both groups, manufacturer financing was less desirable while the respondents split on the desirability of a FPL financing option across equipment types.

Exhibit 3
Average Coefficient Estimates From Conjoint Logit Model

	PRICE	REBATE	SAVINGS	FPLFIN	MFGFIN
CAC	-0.012	0.018	0.028	1.47	-1.06
Heat Pumps	-0.010	0.006	0.034	-1.72	-1.67

Along with the coefficients and attribute levels associated with each choice, the observed portion of the utility function ($\beta' X_{ij}$) for each choice also contains a constant that is added in after the estimation. This constant is intended to capture all of the unobserved factors that are likely to influence equipment choice. The equipment choice probabilities are calculated for each individual and averaged across the sample. Predicted probabilities are then compared with actual equipment choice distributions for each demographic segment within FPL's service population. The predicted probabilities are adjusted to match the observed market distribution by adjusting the constant for each equipment choice. The equipment choice distribution is compared to actual 1997 purchase behavior both in and outside the HVAC program. The model distribution is calibrated to the observed SEER distribution in the market to within 1 percent accuracy.

Model Structure

The HVAC Model allows the analyst to change the model inputs to take into account changes in market conditions and alternative program configurations. Market related inputs that can be changed include housing starts and demolition rates as well as equipment costs for both CAC's and heat pumps. Program inputs include program awareness, administration costs, eligibility requirements, and rebate amounts by SEER. The analyst can also vary the time period for which the model is run between 1997 and 2009.

From FPL's customer population, the potential market for HVAC replacements is determined. This is done by dropping those customers who do not own their residences, do not have an existing HVAC system to be replaced, or do not have existing ducts. These customers are dropped as they are either are not eligible to replace since they do not have existing equipment or else are deemed unlikely to undertake the expenditure of replacing their cooling system because they are renters. In the HVAC Model, the FPL Population is divided into 25 demographic segments based on housing type (Single Family Detached (SFD), Single Family Attached (SFA), Mobile Home (MH)), monthly electricity use (Low, Medium, or High for SFD only), and five regions within the service territory.

Once the potential market is determined, the number of annual HVAC replacements is estimated. For each demographic segment the number of HVAC replacements is estimated based on the age distribution of existing equipment within that segment. These annual replacements reflect equipment purchases made both in and outside FPL's HVAC Program.

Once the number of replacements is estimated for the population, the number of replacements is multiplied by the equipment choice probabilities calculated from equation (3) using the logit estimation results. This is done separately for each of the eleven CAC and heat pump options. Multiplying the equipment choice probabilities by the annual number of purchases yields a distribution of purchases across all possible equipment choices both in and outside the program.

The HVAC Model uses 1997 as the base year, and provides predictions up to the year 2009. Given the calibrated equipment choice probabilities and the FPL customer population in 1997, HVAC replacements are estimated for subsequent years by adjusting the population and potential market estimates to account for housing starts, demolition, and changing ages of existing HVAC equipment. In addition, the 1997 HVAC program standards stipulate that past program participants are not eligible to replace equipment through the program. These customers are also subtracted from the potential market for each year. In this manner, the model provides annual purchase predictions for subsequent years taking into account changes in the customer population and potential equipment market.

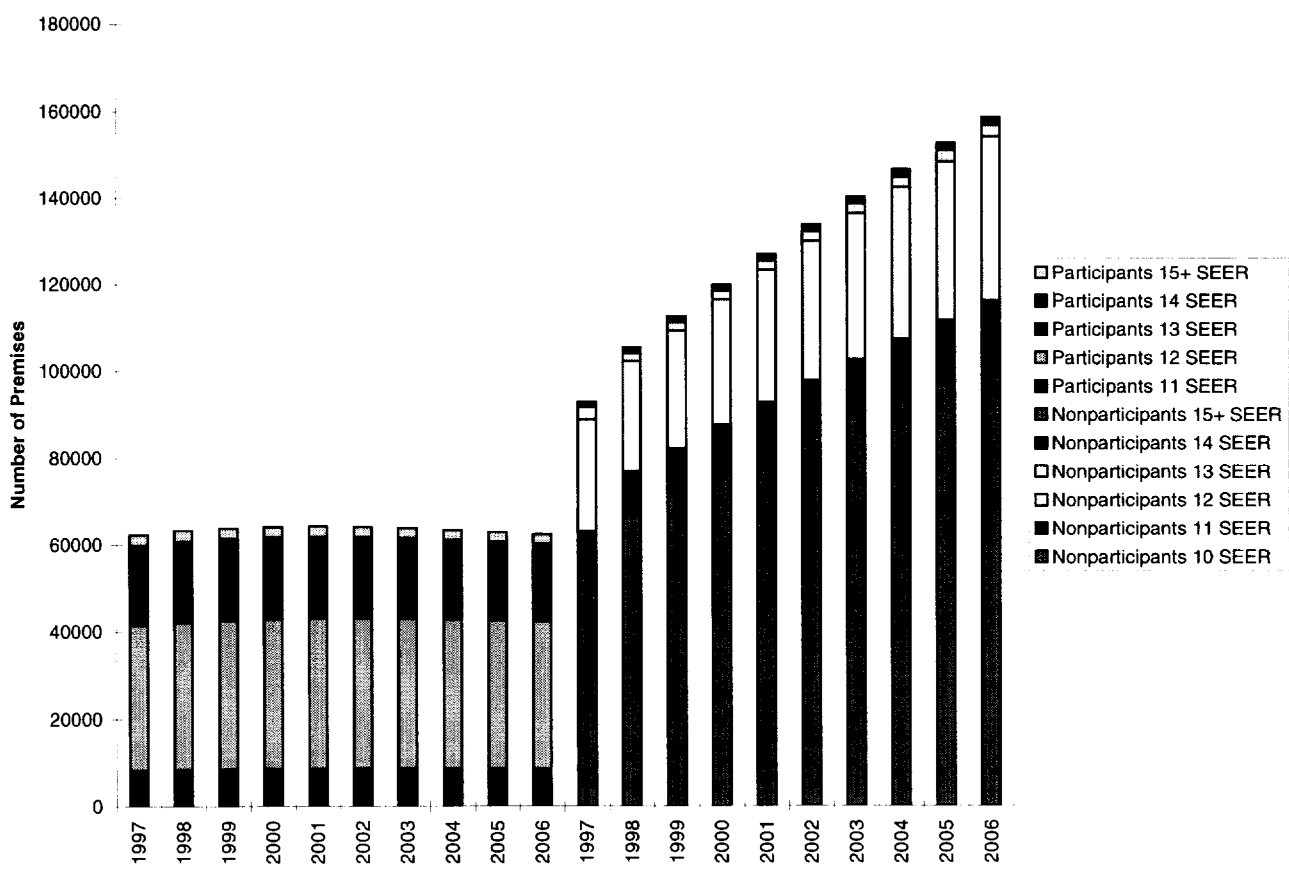
Model Results

Each time the HVAC Model is run, the model estimation results are output to an Excel workbook. The output workbook contains predictions on annual equipment purchases by SEER and demographic segment. In addition, the output workbook contains information on total impacts, dollars spent per kW impact within the program, and a comparison of predicted impacts to FPL impact goals.

The Base Case Scenario using the 1997 program standards, equipment costs, and program awareness levels is shown in Exhibit 4. In this scenario, total HVAC replacements are expected to increase each year from 1997 to 2006, while program participation is expected to remain relatively constant. This is due to the eligibility requirements for the HVAC program, where past participants are ineligible to participate. As a result, while replacements increase every year, the amount of program-eligible purchases remains relatively constant over time, as does program participation.

Annual purchases for the Base Case Scenario are also broken down by SEER in Exhibit 4. 12 SEER units are the favored equipment choice in the program, comprising 54 percent of purchases. This is followed by 14 SEER purchases at 17 percent, and 11 SEER and 13 SEER at approximately 12 percent each. In contrast, purchases outside the program are dominated by 10 SEER units, which make up 50 percent or more of nonprogram purchases each year. As this is the Base Case Scenario, the distribution across SEER levels for program purchases remains unchanged across time as no model inputs have been adjusted for future years.

Exhibit 4
Total Number of Premises Replacing HVAC Systems in a Given Year
By Program Participation and SEER



Cost-Effectiveness and Program Impacts

From a planning perspective, an important output of the HVAC Model is a comparison between incentive dollars spent and MW impact attributable to the HVAC program. This is shown for the Base Case Scenario in Exhibit 5. The cost-effectiveness of the program is shown by comparing the dollars per cumulative kW impact with total MW impacts broken out by building type and usage segment. Cumulative impacts are determined by equipment efficiency and the number of units purchased within the program.

As shown in Exhibit 5, the highest program impacts are found with 12 SEER units, which have cumulative impacts of approximately 25 MW for the SFD Medium Usage and SFA segments. For the SFD High Usage segment, 14 SEER units provide an impact of almost 30 MW, while 12 SEER units provide the highest impact at just under 50 MW. The SFD High Usage segment is also the most cost-effective segment, with about $450 per Summer Peak KW for 11 SEER units and $500 per KW for 12 SEER units.

Exhibit 5
Intersegment Comparison of Cumulative 1997-2006 Summer Peak Demand Impacts And Incentive $/Summer Peak kW

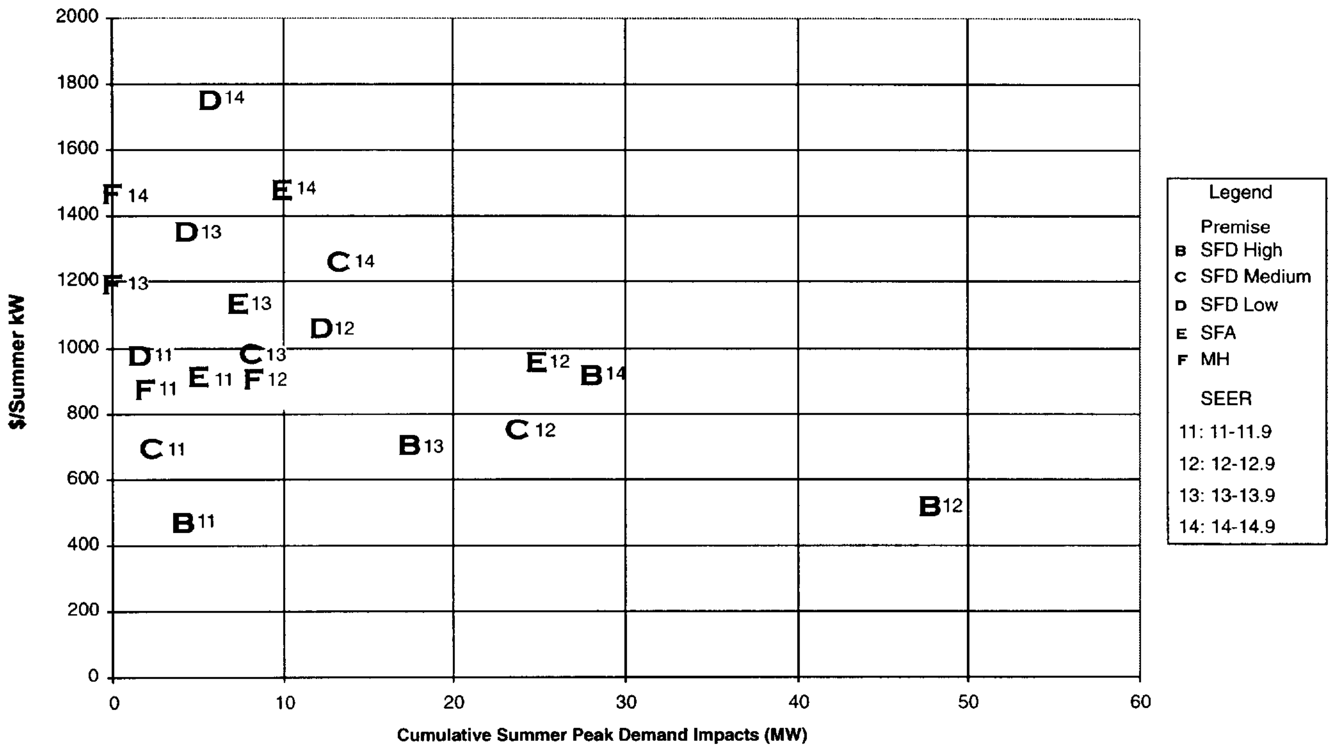
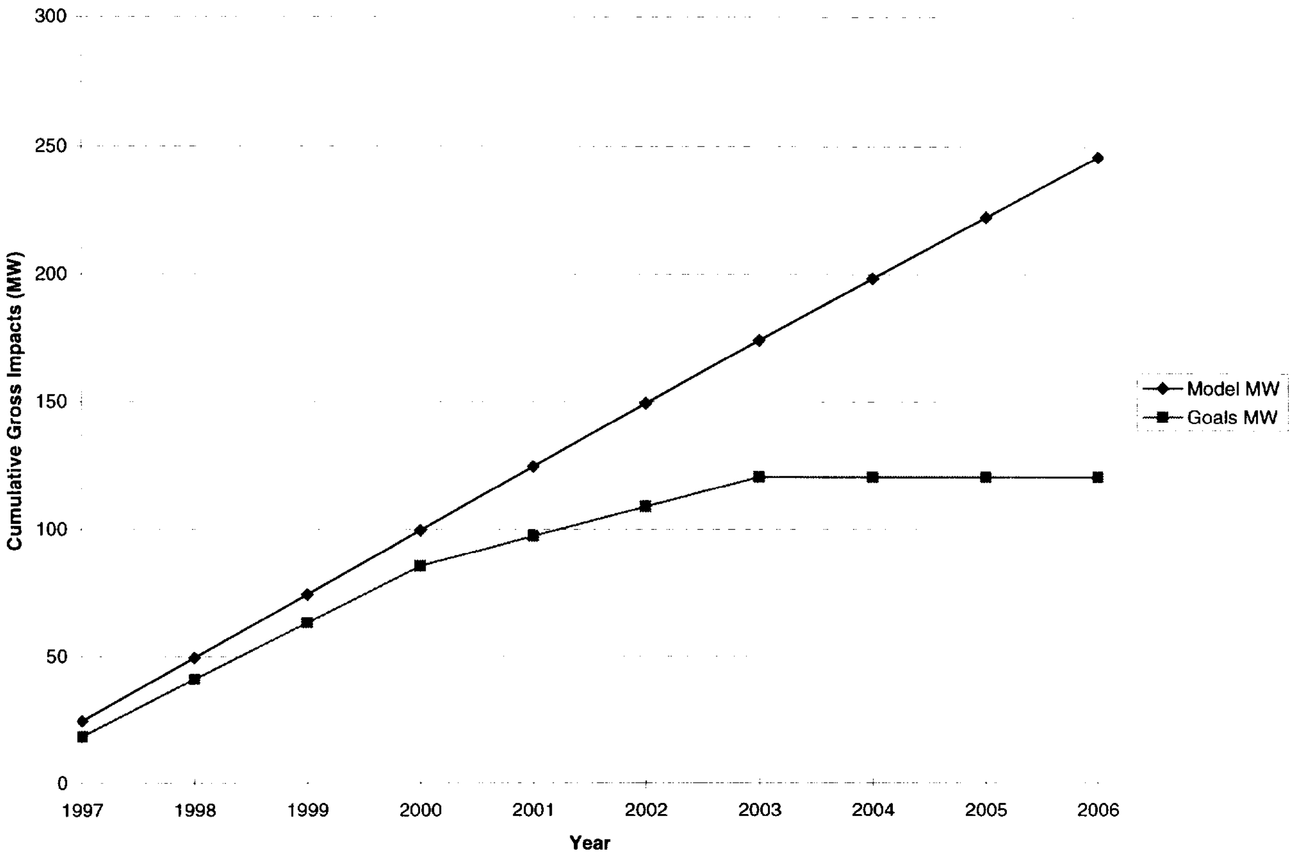


Exhibit 6 shows the cumulative impacts of the Base Case compared with FPL's HVAC Program MW goals. Under the current program standards, the Market Penetration Model predicts that the FPL goal of 121 MW impacts will be met ahead of schedule in 2001. This result combined with the cost effectiveness information output in the Exhibit 4 suggests that there are alternative program configurations that will enable FPL to reach its program goals in a more cost effective manner. How the model can be used to address this issue is demonstrated in the following section.

Exhibit 6
Comparison of Model Impact MW Estimates and FPL Impact Goals

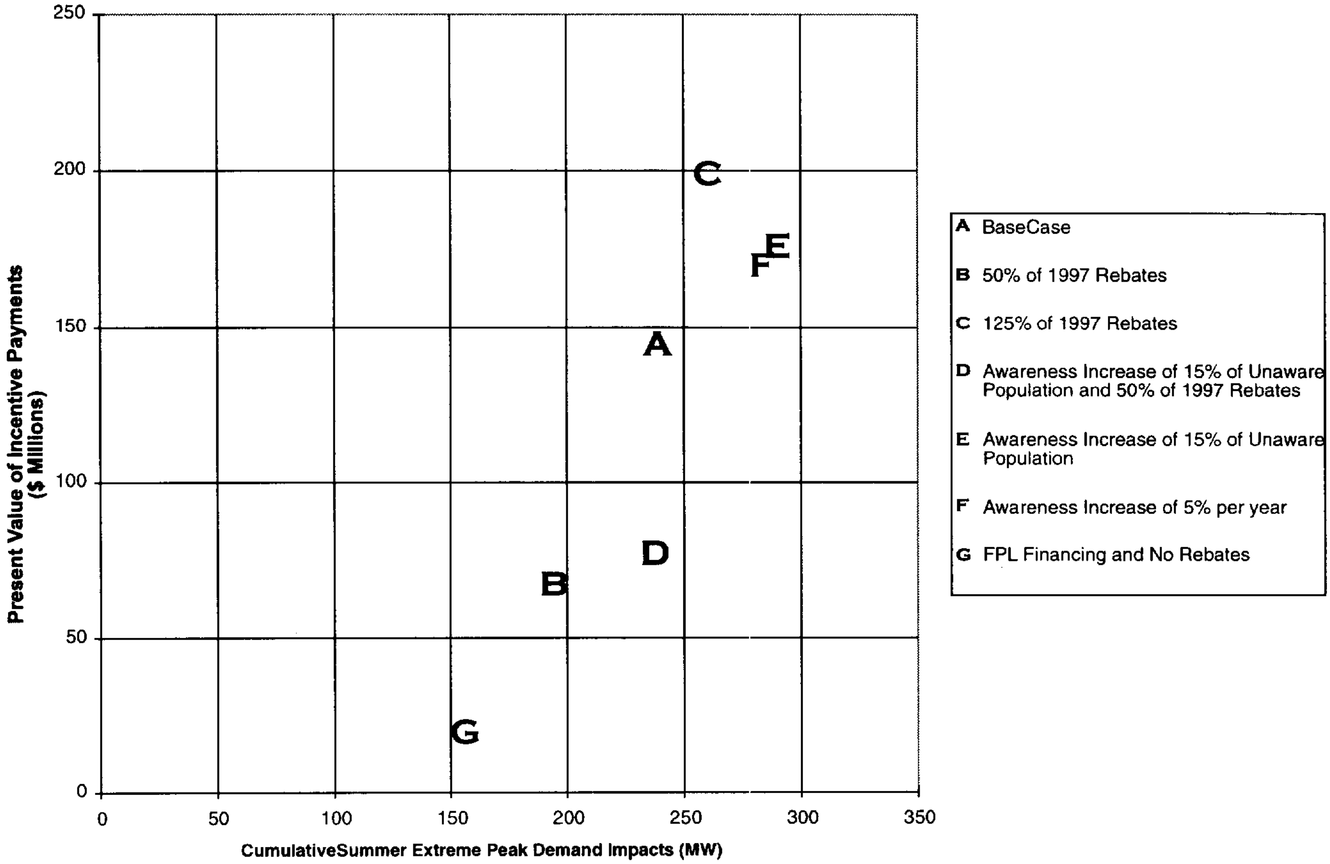


Alternative Program Scenarios

The HVAC Model was developed to evaluate the cost-effectiveness of the current Residential HVAC Program by examining alternative program configurations. Again, controlled conjoint experiments provide a means to estimate how customers will react to program and market conditions that currently do not exist. Once the model has been estimated and the Base Case Scenario established, the HVAC Model inputs can be adjusted to evaluate alternative program configurations.

The Base Case Scenario is compared with alternative program configurations in Exhibit 7. The Base Case Scenario (labeled as A) shows a total impact of just under 250 MW at a cost of approximately $150 million in rebates over the period of 1997 to 2006. Several alternative program configurations are shown for comparison to the Base Case. A program where rebates are reduced by 50 percent for all SEER levels is labeled Scenario B. This reduces incentives paid to $75 million, while cumulative impacts fall 20 percent to approximately 200 MW. The most cost-effective alternative program configurations is Scenario D, with a 50 percent reduction in rebates and a one time increase in program awareness of 15 percent in 1998. The increase in awareness offsets the decrease in participation due to the cut in rebate. The net result is an almost 50 percent reduction in incentives paid relative to the Base Case, with almost no change in impacts over the same time period.

Exhibit 7
Comparison of Cumulative 1997-2006 Summer Extreme Peak Demand Impacts
And Present Value of Incentive Payments
Alternative Program Configurations, HVAC Program



Conclusion

Predicting future market behavior remains an uncertain business. Conjoint analysis, a powerful stated preference technique, provides a useful tool in helping predict future behavior under changing market conditions. This paper describes the development of a market penetration model based on data obtained in controlled conjoint analysis experiments. Based on the stated preference data and calibrated to actual purchase behavior both in and outside the HVAC program, the Model provides estimates of HVAC replacements from 1997 up through 2009. The HVAC Model allows key market variables such as equipment costs, program rebates, and program awareness to be modified by the researcher as model inputs. This allows the researcher to predict purchase behavior under alternative market conditions and program scenarios. The ability to predict future behavior under changing market conditions is critical for program planning, and the Model provides a valuable tool for developing a program that meets impact goals in the most cost-effective manner.

<u>References</u>

Christakis, Nicholos A. and Paul D. Allison. "Logit Models for Sets of ranked Items," *Sociological Methodology* 24:199-228.

Louviere, Jordan J. and David A. Hensher. 1983. "Using Discrete Choice Models with Experimental Design Data to Forecast Consumer Demand for a Unique Cultural Event." *Journal of Consumer Research* 10:348-361.

Maddala, G. S. 1983. *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*. New York: Academic Press: 105-142.

Moore, William L. 1980. "Levels of Aggregation in Conjoint Analysis: An Empirical Comparison," *Journal of Market Research* 11:516-523.