EXPERIMENTS IN SURVEYS: INFORMATION TREATMENTS IN EVALUATING A NEW TECHNOLOGY

Daniel H. Hill, The University of Michigan

Since Rousseas and Hart's (1951) work on composite indifference functions, the survey experiment has come to be a standard tool for investigating peoples' preferences and values. While the methodology has been used in studying marketed goods,[1] it is especially useful in investigating non-market phenomenon or extra-market goods.[2] In addition to a shared goal of understanding peoples subjective preferences and values, most applications of survey experiments share one common design element--the subjects are presented with a number of combinations of situation characteristics (i.e. factor values) and are asked, one way or another, to rank the various combinations.

While the empirical evidence is not yet conclusive, those studies which have compared survey experimental data with more objective data find that the former data tend to yield exaggerated effects of treatments on the dependent variable.[3] This tendency toward exaggeration has come to be termed "hypothetical bias", and a major goal of this paper is to suggest experimental design and analytic procedures which will minimize its impact.

For this reason, in the present study, we deviate from tradition by asking respondents about the objective productive utility of a <u>single</u> configuration of characteristics of a virtually unknown new technology--electric vehicles. We use only one combination of factor values per respondent because it is our opinion that, to a large extent, hypothetical bias is the result of respondents trying to impose internal consistency between their choices regarding the various combinations of treatments they are presented. We limit our population of inference to managers of commercial (SIC codes 1-89) vehicle fleets in the United States. Our goal is to estimate the relative importance of two (technically competing) characteristics of such vehicles. These are the total life-cycle costs of the vehicle, and its range (in miles) on a single charge of its batteries.

The experiment was conducted by first educating the respondent about the characteristics of electric vehicles and then asking whether and how many electric vehicles he would be willing to use in his business operations. In the course of "educating" each respondent a different combination of range and cost values is given depending upon which cell of a factorial design matrix the respondent is randomly assigned. Estimates of the relative importance of the two factors are obtained by examination of the cross-sectional difference in numbers of electric vehicles respondents are willing to use, and the associated values of the two treatments.

The paper is organized into three sections and a summary. In the first section, the relative merits of various experimental design options are discussed. In Section 2 we describe the Survey of Commercial Fleet Managers and identify the econometric techniques needed to estimate the parameters of the response surface (i.e. the demand function for commercial electric vehicles). In the final section we present the results of the study and test the effectiveness of the design in eliminating hypothetical bias.

Experimental Design Considerations

There are two ways of assessing the relative importance of the individual characteristics of a new technology such as electric vehicles. The first is to conduct a demonstration program with many prototypical configurations of vehicles. The second is to conduct a survey experiment.

There are advantages and disadvantages of each of these methods. The first method, which has the advantage of providing "hands-on" experience with the technology, has the disadvantage of being enormously expensive. Furthermore, if the prototypes differ in important ways from the actual technology (for instance, because the prototypes are not mass-produced) then the demonstration will yield erroneous estimates of the usefulness of the technology.

The second (or survey experiment) method is much cheaper, provides factual information on such important objective variables as, in the present example, the distribution of daily vehicle usage, as well as information on the responsiveness of a representative sample of potential end-users to the technical characteristics in question (costs and range). Its disadvantage lies in the fact that the technology described is a hypothetical one, and the response <u>itself</u> is hypothetical in the sense that the respondent is not required to actually use it. Thus the survey experiment is subject to the type of hypothetical bias noted in the introduction.

Our desire to minimize hypothetical bias has implications for both the level of factor values and the number of combinations of factors presented to each respondent. In setting the levels of the factor values we must trade-off potential gains in statistical efficiency (which might result from having a wide range of values) for reductions in potential hypothetical bias (which could come about by keeping the factor values within a plausible range). Because we are more concerned with minimizing hypothetical bias than in obtaining precise estimates of the elasticities, we confine the variation in the factor values to relatively narrow bands.

Elsewhere,[4] we have shown that economic theory indicates that the derived demand for electric vehicles relative to conventional vehicles should be a non-linear function of their relative life cycle costs and their range. In order to capture this non-linearity it is necessary to have at least three values of each treatment variable. For these reasons we chose to employ a factorial design in which the life-cycle cost values of electric relative to conventional vehicles of "ten percent less than", "the same as" and "fifteen percent more than" are cross-randomized with range values of "thirty", "sixty" and "ninety" miles per charge.

Just as the level of treatments can effect the salience to the respondent of the hypothetical nature of the question, so too can the number of treatment values he is presented. There are a number of ways in which we could provide variation in the information presented to respondents which would, in conjunction with their responses to questions about the usefulness of electric vehicles, allow us to make estimates of the impacts of range and price on demand. One method which is often used in social-psychology is the vignette or factorial survey methodology, in which a single respondent is sequentially presented with a number of combinations of the various treatments (e.g. range and price) and after being presented with each, asked to provide a response.[5] While this technique provides more 'observations' per interview, it has the potential of introducing biases to the response estimation procedure for two reasons.

First, the respondent will quickly realize that "some sort of game is being played" in the interview, with respect to the values of the treatment variables, and is likely to alter his responses to the questions in such a manner as not to appear inconsistent, and thereby stupid.

Second, it will increase the extent to which the technology being discussed is perceived as hypothetical rather than actual. The first time the interviewer presents the respondent with a set of technical characteristics he will tend to take them at face value (unless, of course, he knows a lot about the technology prior to the interview and the characteristics are inconsistent with this knowledge). The moment the respondent is given a new set of values of the treatments, he will realize that the technology is not 'set in concrete', and may wonder if it really exists at all.

For these reasons, and for reasons of costs and interview length, we decided not to follow a vignette design, but rather present each respondent one and only one set of technical characteristics, and ask the questions of usefulness only once per respondent.

Data and Econometrics

In September of 1983, the Survey Research Center began interviewing a national probability sample of nearly six-hundred commercial vehicle fleet managers. The bulk of the survey concerned obtaining information on such objective factors as numbers of vehicles used in business applications, daily mileage, and vehicle assignment and replacement criteria.

At the end of the interview we told the respondent what he could reasonably expect from an electric vehicle and then asked if such a vehicle would be useful in his business. If the respondent said that it would we asked how many such vehicles he would be willing to use. Six pieces of information (reliability, recharge time, top speed, capacity, range, and life-cycle costs relative to conventional vehicles) were provided. The experimental treatments consisted of varying what the respondent was told about the relative costs and the range. The respondent was told the life-cycle cost of an electric vehicle was either ten percent less than, the same as, or fifteen percent higher than a conventional vehicle and that it had a range of either thirty, sixty, or ninety miles between recharges. In all there were nine cells in the design matrix, and respondents were randomly assigned to one of these cells.[6] The design matrix (with, for each cell, the final sample size, proportion of respondents saying an electric vehicle with the given characteristic would fit into their operations, and the number of such vehicles they could use) is presented in Table 1 below.

Table 1

The Design Matrix

	RANGE		
Life-Cycle Costs	30 miles	60 miles	90 miles
10% less than conventional	65 .292 1.88	42 .429 1.77	48 .500 1.43
Same as conventional	60 .383 1.52	45 .356 2.00	53 .396 1.33
15% more than conventional	52 .154 1.60	58 .345 1.55	57 .351 1.80

The reader will note that in estimating the price and range effects on the number of electric vehicles variable we have a classic example of a limited (censored) dependent variable. The use of ordinary least squares to estimate these responses would result in underestimates of the price and range responses because it does not correct for the fact that at some prices and ranges some managers not only do not want electric vehicles, but would dispose of them if they had them. Fortunately, the proportion of managers saying that an electric vehicle of the sort described would fit into their fleet operations is sufficiently close to .5 that we need not worry about heteroscedasticity and can therefore use a standard Tobit estimation procedure. That is, following Tobin's (1958) parameterization:

$$hEV = \beta' X_i + \mu \quad \text{if } \#EV>0$$

$$hEV = 0 \quad \text{otherwise}$$

where:

h is the inverse of the standard deviation of the dependent variable

EV is the natural logarithm of the number of electric vehicles the respondent said he would use,

β is the vector of standardized cost and range elasticities,

X_i is the vector of independent variables including the natural logarithms of life-cycle costs and range, and

μ_i is a homoscedastic normally distributed error term with zero mean and unit variance.

The log-likelihood function for this model can be written as:

$$\log L = \sum_0 \log F(-\beta' X_i) + N_1 \log(h) - \frac{1}{2}\sum_1 (hEV_i - \beta' X_i)^2$$

where F is the cumulative normal density function.

We choose the double logarithm specification because of its ease of interpretation, and because, as Houthakker (1961) has argued, of its yielding a better fit to a wider range of data than alternative specifications.

Although we have gone to considerable lengths to design the experiment in such a manner as to minimize the possibility of hypothetical bias, there is no guarantee that these measures have been successful. Even if we have not removed all the hypothetical bias, there may be ways of interpreting the results which yield meaningful implications. Juster and Shay (1964) have argued that the extent of exaggeration from hypothetical data can be expected to be the same for each of the β's. If this is so, then, although the level of the individual β's may be biased, their ratios will not be. Fortunately for the particular problem at hand, it is just such a ratio which is of importance to the designers of electric vehicles.

To see this, note that the derived demand equation presented above can (suppressing the individual subscript i) be written as:

$$\ln(\#EV) = a + \beta_1 \ln(C(R)) + \beta_2 \ln(R) + \mu.$$

where R is the design range, β_1 is the cost elasticity, β_2 is the range elasticity, and C(R) is an engineering relation of vehicle cost to range.[7]

An expression for the design range which would maximize the demand for electric vehicles in the commercial sector can be obtained by maximizing the above function with respect to R. The first order condition for a maximum is that the first derivative of (9) equal zero. That is:

$$(10) \qquad \frac{\partial \ln(\#EV)}{\partial R} = \frac{\beta_1}{C(R)} \frac{\partial C(R)}{\partial R} + \frac{\beta_2}{R} = 0$$

Subtracting the second term of the right-hand side from both sides of the equation, dividing through by β_1, and multiplying by R yields:

$$(11) \qquad \frac{R}{C(R)} \frac{\partial C(R)}{\partial R} = -\frac{\beta_2}{\beta_1}$$

Which means that the optimal range is determined by the characteristics of the cost function C(R) and the ratio of the range and cost elasticities. If Juster-Shay hypothesis is correct, then the observed response $\hat{\beta}_i$ is the product of the actual ($\tilde{\beta}_i$) and some constant (λ) greater than one--an "exaggeration constant" which is canceled out by the division on the right-hand side of the last expression.

Finally, while a major advantage of randomized assignment of cases to treatments is that one need not control for other factors in order to get unbiased estimates of the response (for members of the finite population from which the sample was drawn), we can learn more by specifying the model more completely. Explicitly, we include measures of characteristics of the firm that might influence a manager's decision about the number of electric vehicles to say he could use. These include the (natural logs of) a) number of vehicles currently in the fleet, b) proportion of the fleet which are passenger cars, c) proportion of the fleet which are trucks, d) proportion of the fleet made up by vans, e) number of passenger cars typically driven less than sixty miles per day, and f) the number of trucks and vans typically driven less than sixty miles per day.

Evaluation of Results

The Tobit regression coefficients for the derived demand model are presented in column one of Table 2. The estimated life-cycle cost elasticity of demand is negative, significant, and more than four times as large in absolute value as the (positive and significant) range elasticity. The implications of this pattern of results is that if the developers of electric vehicles wish to maximize the usage of electric vehicles in the commercial sector, then they should strive to configure a vehicle which features economy even at the expense of range.

In addition to the price and range elasticities Table 2 presents a number of effects of firm specific characteristics which are of interest. Although it is not quite large enough to attain statistical significance at conventional levels of confidence, the effect of fleet size is relatively large and positive. Its size indicates that for each ten per cent increase in fleet size the demand for electric vehicles increases by about one and a half percent. This finding is consistent with both theory and common sense. The larger the fleet the greater the chance that at least one existing vehicle is used in a manner which is consistent with electric vehicle technology. It is also interesting to note that the number of passenger cars currently in the fleet which travel sixty or fewer miles per day has a positive, significant effect on the level of demand for electric vehicles. Finally, there is a strong (significant and persistent) positive impact of (the natural logarithm of) the variable measuring the proportion of the firms current fleet which is made up of vans. It is often the case that vans are used in commercial applications as 'mobile toolboxes'. They are driven by service technicians to perform service calls, often locally. Such applications are well within the scope of near-term electric vehicles. The implication of this last finding is that if a potential manufacturer had to decide on just one body type of electric vehicle, he should choose to produce a van.[8]

While a direct test of the effectiveness of the design measures we employed to eliminate hypothetical bias is not possible, some indication of their effectiveness can be obtained by examining the effects of interview specific variables on the dependent variable. If the survey experiment is truly measuring the parameters of the derived demand for electric vehicles, then it should not matter which interviewer is asking the question.

Columns two and three of the Table 2 present the Tobit regression results we obtain when controls are made for interviewer effects by inclusion of interviewer identification dummies, or measures of interviewer characteristics. An examination of the values of the log-likelihood function indicates that the inclusion of the interviewer controls does significantly affect the fit of the model. The likelihood ratio test of the null hypothesis that interviewer identification dummies do not affect the expressed demand for electric vehicles yields a χ-square of 22 with thirteen degrees of freedom. What is more, as a comparison of the log-likelihood values for the models presented in columns two and three will reveal, the interviewer effects operate through something other than the objectively measured interviewer characteristics included in the later model. While interviewer efficiency does have a significant effect on the expressed demand for electric vehicles, the set of interviewer dummies 'picks up' significantly more variation in the dependent variable than does the set of interviewer characteristics.

Finally, while the absolute level of estimated effects of the treatment variables are noticeably affected by the inclusion of interviewer controls, their relative importance is not. In all three versions of the derived demand model, the importance of life-cycle costs is somewhat more than four and a quarter times

Table 2

TOBIT REGRESSION COEFFICIENTS OF THE COMMERCIAL DEMAND FOR ELECTRIC VEHICLES (Dependent Variable = ln (#EVs))

Variable	Model 1	Model 2	Model 3
ln (cost)	-1.86* (-2.50)	-2.23** (-2.98)	-2.15** (-2.87)
ln (range)	.45** (2.79)	.49** (3.01)	.48** (2.98)
ln (fleet size)	.17 (1.50)	.19 (1.77)	.17 (1.45)
ln (#cars < 60 m.p.d.)	.53** (3.97)	.57** (4.34)	.54** (4.10)
ln (#trucks < 60 m.p.d.)	.08 (.64)	.15 (1.21)	.13 (1.12)
ln (proportion cars)	-.05 (-1.00)	-.05 (-1.07)	-.05 (-1.01)
ln (proportion trucks)	.06 (1.24)	.08 (1.87)	.06 (1.32)
ln (proportion vans)	.18** (3.97)	.20** (4.50)	.17** (3.85)
Interviewer ID Dummies	-	‡	-
Interviewer Characteristics:			
Efficiency	-	-	.19* (2.45)
SRC Experience	-	-	-.02 (-.45)
Sex	-	-	-.34 (-1.56)
Age	-	-	-.01 (-1.44)
Pilot Study	-	-	.06 (.37)
Log-Likelihood Value	-430.5	-417.7	-425.5
Pseudo R^2	18.4%	19.8%	18.7%
n	474	474	474

*Significant at the 95% level of confidence.

**Significant at the 99% level of confidence.

‡This set of 13 dummy variables is significant at the 95% level of confidence.

that of range. This same ratio of effects is found to hold in a wide variety of alternative specifications of the derived demand equation including the non-linear constant elasticity of substitution (CES) as well as the log-linear CES with random effects.[9] These findings lend credence to the Juster-Shay hypothesis that hypothetical bias effects all factors equally.

Summary

In the present paper we have suggested two design options which were intended to minimize the possibility of hypothetical bias in survey experimental data. These were the limiting of treatment values to a narrow range of plausible values, and presenting the respondent with only one combination of factor values. We incorporated these options in the design of a survey experiment to estimate the importance of life-cycle costs and vehicle range in determining the productive utility of electric vehicles in commercial applications. The fact that interviewer-specific variables had a significant impact on the goodness of fit of the derived demand estimates from our experimental data suggests that hypothetical bias may not have been eliminated by the experimental design. Nevertheless, the relative importance of the two treatment variables remained unaffected by the inclusion of interview variables (or, for that matter, by employing radically different specifications of the demand model). This last finding is consistent with the Juster-Shay hypothesis that hypothetical bias affects all estimates equally--in which case the substantive implication of the experiment is that developers of electric vehicles should strive to produce an economical short-range vehicle if they wish to obtain an appreciable share of the commercial vehicle market.

[1]Juster and Shay (1964) employed a survey experiment to obtain estimates of interest and loan period elasticities of demand for private borrowing.

[2]Social psychologists use the technique (termed vignette methodology or factorial survey in their field) to investigate such non-market phenomena as peoples' perceptions regarding the seriousness of specific crimes and the justice of punishments (See Rossi and Nock 1982), while natural resource economists use it to estimate the demand for public goods (see Schulz, d'Arge, and Brookshire, 1981).

[3]Juster and Shay (1964) found the interest rate elasticity of private borrowing to be significantly stronger with survey experimental data than with actual time-series market data. Similarly, Bishop and Heberlein (1979) found the estimated willingness to buy hunting permits from a survey experiment was greater than from an experiment in which permits were actually bought from hunters.

[4]See Hill (1984).

[5]For a description of these techniques see Rossi and Nock (1982).

[6] Randomization was accomplished at the initial sample selection stage by creating a selection variable which ranged from 1 to 9. The interview was conducted using SRC's Computer Assisted Telephone Interviewing facility (CATI). The CATI system works by displaying the question the interviewer is to read on a CRT screen. The interviewer enters the respondent's answer on the keyboard and the CATI system displays the next logical question. The wording of the question which contains the information on electric vehicle characteristics was automatically varied depending on the value of the selection variable.

[7]This function is increasing in R because the only way of increasing the range of an electric vehicle is to add more (or more expensive) batteries. This increases both initial costs and operating costs because the batteries are expensive and heavy, and the amount of energy required to propel the vehicle increases with vehicle weight.

[8]It may not be a coincidence that the only mass produced electric over-the-road vehicle in the world is the Lucas Electric Van produced in Bedford England.

[9]See Hill (1984).

References

Berg, Mark R., Converse, Muriel J., and Hill, Daniel H., Electric Vehicles in Commercial Sector Applications: A Study of Market Potential and Vehicle Requirements, (Ann Arbor Michigan: Institute for Social Research May 1984).

Bishop, Richard C., and Heberlein, Thomas A., "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?", American Journal of Agricultural Economics, 61, (1979) pp. 926-930.

Hill, Daniel H., "The Derived Demand for Not-Yet-Existing Inputs", (Ann Arbor, Mich: The Survey Research Center, July 1984).

Juster, F. Thomas, and Shay, Robert P.,Consumer Sensitivity to Finance Rates: An Experimental and Analytical Investigation, (New York: NBER Occasional Paper no.88, 1964).

Rossi, Peter H. and Nock, Steven L. (eds.) Measuring Social Judgements,(Beverly Hills, Calif: Sage Publications, 1982) pp. 15-68.

Rousseas, Stephen W. and Hart, Albert G., "Experimental Verification of a Composite Indifference Map", Journal of Political Economy 59 (August 1951) pp. 288-318.

Schulze, W.D., d'Agre, R.C., and Brookshire, D.S. "Valuing Environmental Commodities: Some Recent Experiments", Land Economics 57 (1981) pp. 151-172.

Tobin, James, "Estimation of Relationships for Limited Dependent Variables", Econometrica 26 (1958) pp. 24-36.