The Navigation Economic Technologies Program

November 2007

NETS

navigation · economics · technologies

MONTE CARLO ANALYSIS OF SP-OFF-RP DATA

US Army Corps of Engineers®

IWR Report 07-NETS-P-03

Navigation Economic Technologies

The purpose of the Navigation Economic Technologies (NETS) research program is to develop a standardized and defensible suite of economic tools for navigation improvement evaluation. NETS addresses specific navigation economic evaluation and modeling issues that have been raised inside and outside the Corps and is responsive to our commitment to develop and use peer-reviewed tools, techniques and procedures as expressed in the Civil Works strategic plan. The new tools and techniques developed by the NETS research program are to be based on 1) reviews of economic theory, 2) current practices across the Corps (and elsewhere), 3) data needs and availability, and 4) peer recommendations.

The NETS research program has two focus points: expansion of the body of knowledge about the economics underlying uses of the waterways; and creation of a toolbox of practical planning models, methods and techniques that can be applied to a variety of situations.

Expanding the Body of Knowledge

NETS will strive to expand the available body of knowledge about core concepts underlying navigation economic models through the development of scientific papers and reports. For example, NETS will explore how the economic benefits of building new navigation projects are affected by market conditions and/or changes in shipper behaviors, particularly decisions to switch to non-water modes of transportation. The results of such studies will help Corps planners determine whether their economic models are based on realistic premises.

Creating a Planning Toolbox

The NETS research program will develop a series of practical tools and techniques that can be used by Corps navigation planners. The centerpiece of these efforts will be a suite of simulation models. The suite will include models for forecasting international and domestic traffic flows and how they may change with project improvements. It will also include a regional traffic routing model that identifies the annual quantities from each origin and the routes used to satisfy the forecasted demand at each destination. Finally, the suite will include a microscopic event model that generates and routes individual shipments through a system from commodity origin to destination to evaluate non-structural and reliability based measures.

This suite of economic models will enable Corps planners across the country to develop consistent, accurate, useful and comparable analyses regarding the likely impact of changes to navigation infrastructure or systems.

NETS research has been accomplished by a team of academicians, contractors and Corps employees in consultation with other Federal agencies, including the US DOT and USDA; and the Corps Planning Centers of Expertise for Inland and Deep Draft Navigation.

For further information on the NETS research program, please contact:

Mr. Keith Hofseth
NETS Technical Director
703-428-6468

Dr. John Singley
NETS Program Manager
703-428-6219

U.S. Department of the Army
Corps of Engineers
Institute for Water Resources
Casey Building, 7701 Telegraph Road
Alexandria, VA 22315-3868

The NETS program was overseen by Mr. Robert Pietrowsky, Director of the Institute for Water Resources.

November 2007

navigation · economics · technologies

MONTE CARLO ANALYSIS OF SP-OFF-RP DATA

Prepared by:

Kenneth E Train, PhD
University of California, Berkley

Wesley W Wilson, PhD
University of Oregon

For the:

Institute for Water Resources
U.S. Army Corps of Engineers
Alexandria, Virginia

IWR Report 07-NETS-P-03

www.corpsnets.us

Table of Contents

Tables

Monte Carlo Analysis of SP-off-RP Data*

by

Kenneth E. Train†
University of California, Berkeley
and
Welsey W. Wilson‡
University of Oregon

November 9, 2007

Abstract

Sp-off-rp data is a recent innovation in choice modeling. Revealed data reflect choices made in a real-world setting. Stated preference data are then constructed by asking the decision-maker whether he or she would choose the same alternative or switch to another alternative if the attributes of the chosen alternative were less desirable in ways specified by the researcher and/or the attributes of nonchosen alternatives were more desirable in specified ways. This construction, called "stated-preference off revealed-preference" (sp-off-rp), can increase the realism of the stated-preference task, relative to standard sp exercises, but creates endogeneity. In this paper, we present a series of Monte Carlo exercises that explore estimation on this type of data, using an estimator that accounts for the endogeneity. The results suggest that sp-off-rp data can provide large efficiency improvements over the use of revealed-preference data alone, especially in small samples. The results also suggest that the sp-off-rp design yields substantial efficiency gains over standard sp designs when, as expected, respondents are able to answer sp-off-rp questions more accurately than standard sp questions due to their greater realism.

JEL CODES: C25 - Discrete Regression and Qualitative Choice Models ; C42 - Survey Methods; C51 - Model Construction and Estimation; C81 - Methodology for Collecting, Estimating, and Organizing Microeconomic Data

---

* The authors gratefully acknowledge research support from the Navigation and Economics Technology Program of the Institute for Water Resources.

† Department of Economics, University of California, Berkeley CA 94720-3880, voice: 415-291-1023, fax: 415-291-1020, train@econ.berkeley.edu.

‡ Department of Economics, University of Oregon, Eugene OR 97405, voice: 541-346-4690, fax: 541-346-1243, wwilson@uoregon.edu

1. Introduction

Consumers' preferences are often estimated by supplementing data on choices that consumers have made in market settings, called "revealed-preference" (rp) data, with data on choices that consumers say they would make, called "stated-preference" (sp) data. In a typical sp experiment, the researcher constructs hypothetical choice situations, each of which consists of two or more alternatives among which the respondent is asked to choose. The attributes of the alternatives are varied over experiments to provide the variation needed for estimation of underlying preference parameters. The purpose of these sp experiments is to generate variation in attributes when the attributes in the market conditions that produce the rp data exhibit insufficient independent variation to allow precise estimation. Examples include Ben-Akiva and Morikawa (1990), Hensher and Bradley (1993) and Hensher et al. (1999) within a logit specification and Bhat and Castelar (2001) and Brownstone et al. (2000) using mixed logit.

"Pivoting" has been used by some researchers to enhance the realism of sp experiments, by constructing alternatives for sp experiments that are similar to ("pivoted off") an alternative that the agent chose in a market setting. For example, in examining route choice, Rose et al. (forthcoming) asked each respondent to describe a recent trip. Hypothetical routes were designed with times and costs constructed as some percent above or below those of the recent trip. The respondent is then asked to choose among these hypothetical routes. The recent trip with its observed times and cost is either included or excluded from the sp choice set, depending on the design of the experiments. Other applications include Hensher (2004), Caussade et al. (2005), Hensher and Rose (2007), and Greene et al. (2006).

Fowkes and Shinghal (2002) and Train and Wilson (2007) have proposed and implemented an alternative way of constructing sp experiments that has the potential to be more effective in eliciting preferences, while also being more realistic for the respondent than either standard or pivoted sp experiments. The respondent's choice in an rp setting is observed and then the respondent is asked which of the rp alternatives he or she would choose if the attributes of the chosen alternative were made worse and/or the attributes of any of the unchosen alternatives were made better. Take, for example, a mode choice situation in which a respondent has chosen bus when car, bus, and rail are available for the commute to work. The respondent is then asked such questions as: ``Would you have chosen bus if the bus fare were $1.50 instead of $1.00?" or ``Would you have switched to rail if the trains were 10 minutes faster than they are now?" In these questions, the respondent faces the same alternatives as in the rp setting except for a specified change in one or more of the attributes.

A distinguishing feature of these questions is that they incorporate the fact that a change in the respondent's rp choice can occur only if the attributes of the chosen alternative are made worse or the attributes of the non-chosen alternatives are improved. By determining the extent to which the attributes of the chosen alternative must be worsened, or the attributes of non-chosen alternatives improved, in order to induce the respondent to change, the underlying preferences of the respondent are revealed.

Train and Wilson (2007) call this procedure “sp-off-rp” because the stated-preference questions are created from the respondent’s revealed-preference choice. Sp-off-rp questions can be considered a form of pivoting; however, they differ from the usual pivoted designs in two important ways. First, with the usual pivoted designs, the respondent faces whatever number of alternatives the researcher constructs and presents to the respondent in the sp task, whereas in sp-off-rp questions the respondent faces the same number of alternatives in the sp task as in the rp task. Second, and related to the first, in sp-off-rp questions, there is a one-to-one correspondence of the sp alternatives to the rp alternatives, whereas in the pivoted experiments cited above each of the sp alternatives corresponds to either one rp alternative (the chosen one) or no specific rp alternative.

Sp-off-rp questions provide several potential advantages relative to standard or pivoted sp designs. First, sp-off-rp questions contain a realism that might not be attained by either standard or pivoted sp experiments, since respondents face the same choice situation, with the same alternatives, in the sp-off-rp questions as they faced in the rp setting. This realism can make respondents more able to accurately assess their choices in the sp-off-rp setting. It can also induce respondents to consider the task thoughtfully since the questions seem meaningful. Second, in standard sp and pivoted sp experiments, the issue necessarily arises of how the respondent assesses or considers the attributes that are not listed in the experiments. For example, in a standard sp experiment for mode choice, the time and cost of the alternatives might be listed, while factors such as risk of delay, the extent of crowding on the bus, whether an easy parking place can be found for the car, etc., are perhaps not included. Inevitably, some attributes are not listed, and it is not clear how the respondent evaluates these non-listed attributes. With sp-off-rp questions, the respondent is asked to consider a change in observed attributes in the rp setting that the respondent faces. The unobserved attributes are therefore, by construction, the same as in the rp setting. This commonality of unobserved attributes across the rp and sp-off-rp data can be explicitly represented and tested in the estimation procedure. Third, the task of estimation is to determine respondents' tradeoffs among attributes as revealed by their choices among alternatives with different attributes. This task is readily served by changing attributes in the directions that are needed to induce a change. Improving an attribute of the chosen alternative cannot change a person's choice and, hence, does not reveal anything about their preferences; neither does worsening the attributes of unchosen alternatives. In standard and the usual pivoted sp experiments, respondents can face choices that reveal little or no information beyond that revealed in their rp choices, since the tradeoffs implied by the rp choice are not taken into consideration in the sp design. In sp-off-rp questions, the attributes are changed in the direction necessary to elicit preference revelation. No matter what the respondent answers in response to these changes, information about preferences is obtained, namely, that the value of the change is either greater than or less than the difference in original utilities.

The potential advantages of the procedure come at an econometric cost. As Bradley and Daly (2000) pointed out, the procedure creates endogeneity in the attributes in the sp-off-rp questions, since these attributes are constructed from the respondent's chosen alternative in the rp setting. Unobserved factors in the rp environment affect the respondent's rp choice and, thereby, affect the attributes in the sp-off-rp setting (since these attributes depend on the rp choice.) As discussed above, the unobserved factors in the rp setting carry forward to the sp-off-rp setting. The sp-off-rp attributes are therefore not independent of the unobserved factors, as usually assumed, but rather depend explicitly upon them. This dependence, if ignored, creates inconsistency in the estimator, as Bradley and Daly (2000) described and documented.

Train and Wilson (2007) developed an econometric method that accounts for this endogeneity and provides a consistent and efficient estimator for sp-off-rp data. They applied the method to data from a survey of shippers, using rp data on the shippers' chosen mode and destination, along with sp-off-rp data on whether the shippers' choices would change if the attributes of the chosen mode/destination became worse. In their application they did not know the true behavioral parameters, and so it was not possible to determine the extent to which the sp-off-rp data provided more precise estimates of them.

In the current paper we use Monte Carlo methods to examine Train and Wilson's econometric procedure for sp-off-rp data. Since the 'true' parameters are known in Monte Carlo data, we are able to assess the extent to which sp-off-rp data increase efficiency, the bias that arises when the endogeneity in the sp-off-rp data is ignored, and the efficiency of sp-off-rp data relative to standard sp data under different assumptions about the error in each type of data. The findings can be summarized as follows:

- For a sample size of 1000 and the parameters in our base specification, sp-off-rp data reduce standard errors for the relevant parameters by a factor of two relative to rp data alone. This result implies that sp-off-rp data provide as large an efficiency gain as quadrupling sample size (since standard errors are inversely proportional to the square root of sample size.)

- For smaller samples, sp-off-rp data provide an even larger gain in efficiency.

- Ignoring the endogeneity in sp-off-rp data creates significant bias in the estimated parameters.

- The econometric method accounts for the possibility that responses to sp-off-rp questions can be influenced by unobserved factors beyond those than enter the rp choice. These may reflect inattention to the task, the inability to conceptualize the situation, or other quixotic aspects of response. As expected, the efficiency gain from sp-off-rp data rises when the variance of these quixotic errors declines. The same result is obtained, also as expected, for standard sp data.

- When these quixotic errors have the same variance in sp and sp-off-rp data, then the two methods provide about the same degree of efficiency. This result implies that the method that obtains more realistic and less quixotic responses provides the greater efficiency, even after accounting for the potential loss of efficiency that dealing the endogeneity in sp-off-rp data entails. Since the motivation for using sp-off-rp questions is to enhance the realism of the choice situation, this result implies that sp-off-rp data are more efficient than standard sp data if indeed this motivating concept is correct.

In the following section, we describe the econometric method for estimating parameters using sp-off-rp data. In section 3, we describe the specification of the Monte Carlo experiments and their results.

2. Econometrics of sp-off-rp data

We describe a fixed coefficient specification first and then generalize to random coefficients.

2.1 Fixed coefficients

Each agent faces a choice among discrete alternatives in an rp setting. The utility that agent n obtains from alternative j is denoted U_{jn}, which is decomposed into observed and unobserved components:

$$U_{jn} = \beta x_{jn} + \varepsilon_{jn}$$

We assume that ε_{jn} is iid type one extreme value, with the result that the model of the rp choice is a standard logit.

To obtain the sp-off-rp data, the researcher gives the agent a series of choice tasks in which the attributes of the alternatives in the rp setting are changed based on the agent's choice in the rp setting, making the attributes of the chosen alternative worse and/or the attributes of the nonchosen alternatives better. The researcher constructs T choice tasks, each consisting of the same alternatives as in the rp setting but with changed attributes. Let $\tilde{x}^i_{jnt}$ denote the attributes for alternative j in choice task t based on alternative i having been chosen in the rp setting. The utility of each alternative in these choice tasks is assumed to take the form:

$$W_{jnt} = \beta\,\tilde{x}^i_{jnt} + \varepsilon_{jn} + \mu^*_{njt}$$

where μ^*_{njt} is a new error term. Under this specification, the agent assesses the alternatives in each choice task using the same utility coefficients β and same ε_{jn} as in the rp setting, but with an additional error term that reflects, e.g., inattention by the agent, pure randomness in the agent's responses, or other quixotic aspects of the choice task. Importantly, the unobserved factors ε_{jn} that affected the agent's choice in the rp setting carry forward to the choice task, since these unobserved factors are not changed. (The assumption that the same β and ε_{jn} enter the rp and sp-off-rp choices can be tested, but for our discussion we take the specification as given.) Let the new error μ^*_{njt} be iid extreme value with scale 1/λ. A large value of parameter λ indicates that there are few quixotic aspects to the sp-off-rp choices, and the agent chooses essentially the same as in an rp situation under the new attributes. Utility can be equivalently expressed as:

$$W_{jnt} = \lambda\beta\,\tilde{x}^i_{jnt} + \lambda\varepsilon_{jn} + \mu_{njt}$$

where now μ_{njt} is iid extreme value with unit scale. The sp-off-rp choices are, therefore, standard logits with ε_{jn} as an extra explanatory variable. Since the ε_{jn}'s are not observed, these logits must be integrated over the conditional distribution of these rp errors. In particular, the probability of alternative *k* being chosen in choice task *t* given that the agent chose alternative *i* in the rp setting is

$$P^i_{knt} = \int \frac{e^{V_{knt}(\varepsilon_{kn})}}{\sum e^{V_{jnt}(\varepsilon_{jn})}}\, f(\varepsilon_n \,|\, U_{in} > U_{jn}\,\forall j \neq i)\, d\varepsilon_n$$

where $V_{jnt}(\varepsilon_{jn}) = \lambda\beta\,\tilde{x}^i_{jnt} + \lambda\varepsilon_{jn}$ and f is the density of $\varepsilon_n = \langle \varepsilon_{1n},\ldots,\varepsilon_{Jn}\rangle$ conditional on alternative *i* having been chosen in the rp setting. This choice probability is a mixed logit, with mixing over ε_n. It is simulated by taking draws from *f*, calculating the logit formula for each draw, and averaging the results. Train and Wilson (2007) derive the conditional density of ε_n based on earlier work by Anas and Feng (1988) and show how to take draws from it.

Under the assumption that μ_{njt} is independent over choice tasks, the probability of the agent's choices in all *T* tasks is the product of logits for the *T* choices, integrated over the conditional distribution of ε_{jn}. The probability of the rp and sp-and-rp choices, which enters maximum likelihood estimation, is the product of (i) the logit probability of the rp choice and (ii) the mixed logit probability of the sequence of sp-off-rp choices conditional on the rp choice:

$$P_n = \int \prod_t \left[\frac{e^{V_{k_t nt}(\varepsilon_{k_t n})}}{\sum e^{V_{jnt}(\varepsilon_{jn})}} \right] f(\varepsilon_n \,|\, U_{in} > U_{jn} \,\forall j \neq i)\, d\varepsilon_n \bullet \frac{e^{\beta x_{in}}}{\sum e^{\beta x_{jn}}} \qquad (1)$$

2.2 Random coefficients

Utility is the same as above except that β is now random with density $h(\beta)$ with underlying parameters (not given in the notation) denoting, e.g., the mean and variance of β. The probabilities are the same as above, except the formulas are now mixed over the distribution of β. The probability that enters the likelihood function is $PR_n = \int P_n(\beta)\, h(\beta)\, d\beta$ where $P_n(\beta)$ is given by equation (1) with β treated as an argument.

3. Monte Carlo analysis

To explore the properties of estimation with sp-off-rp data, we start with a specification that consists of two alternatives, labelled 1 and 2, with two explanatory variables, labeled x and z. Utility contains an alternative-specific constant (α_i), one variable (z_{in}) with a fixed coefficient, and the other variable (x_{in}) with a random coefficient:

$$U_{in} = \alpha_i + \theta \cdot z_{in} + \beta_n \cdot x_{in} + \varepsilon_{in}, \quad i = 1,2, \quad n = 1, \ldots, N.$$

with

$$\varepsilon_{in} \sim iid\ extreme\ value,$$
$$\beta_n \sim N(\bar{\beta}, \sigma^2)$$

The true parameters are specified to be:

$$\alpha_1 = 1, \quad \theta = 1, \quad \bar{\beta} = 1, \quad \sigma = 0.5\,.$$

Each variable for each alternative is specified to be distributed uniformly between 2 and 4, such that the difference between the two alternatives ranges from -2 to +2 for each of the two variables. (In the sections below, each of these elements of the data generation process is revised to examine their impact on the estimator.)

The agent chooses alternative 1 iff $U_{1n} > U_{2n}$ and otherwise chooses alternative 2. Define $d_n^{\ 1} = 1$ if agent *n* chooses alternative 1, = 0 otherwise; and define $d_n^{\ 2}$ similarly. This choice, and the value of the variables *x* and *z*, are the rp data. We now specify the sp-off-rp data. Only one choice task is given to each agent. If alternative *i* is chosen in the rp setting, the value of x_{in} is lowered by r_n proportion, where r_n is uniformly distributed between 0 and 1. Utility in the sp-off-rp situation becomes

$$W_{in} = \lambda(\alpha_i + \theta \cdot z_{in} + \beta_n \cdot (x_{in} - r_n d_n^{\ i} x_{in}) + \varepsilon_{in}) + \mu_{in}, \quad i = 1, 2 \quad n = 1, \ldots, N.$$

where subscript *t* is omitted since there is only one. The new error μ_{in} is specified to be iid extreme value with unit scale after standardizing for the true scale, which is specified to be $\lambda = 4$. This value of the scale was chosen for our initial specification because it is similar to that estimated by Train and Wilson (2007). The agent chooses alternative 1 iff $W_{1n} > W_{2n}$ and otherwise chooses alternative 2. Note that the attribute $(x_{in} - r_n d_n^{\ i} x_{in})$ is correlated with ε_{in} since the agent's rp choice, $d_n^{\ i}$, depends ε_{in}. This correlation constitutes the endogeneity that arises in s-off-rp data.

Each sample consists of 1000 agents. The sample data are simulated and the parameters are estimated data 100 times. In each estimation, 100 randomized Halton draws are used to simulate the integral over the random coefficient β_n and the conditional errors ε_{1n} and ε_{2n} in W_{1n} and W_{2n}.

The results are summarized in Table 1. The mean estimates are very close to the true parameters, with none of the differences being statistically significant.[4] Also, the standard deviations of the estimates are very similar to the mean standard errors, which imply that the standard errors provide reliable information, on average, about the expected sampling error in the point estimates. With the exception of the scale parameter, the standard deviations of the

[4] Since there are 100 runs, the standard deviation of the mean is one-tenth the standard deviation of the point estimates. The t-statistic for the hypothesis that the mean of the sampling distribution of point estimates of, for example, the intercept is equal to 1.0 (it's true value) is (1-.9940)/(0.0753/10)= 0.80.

standard errors are quite small, indicating that the standard errors for any one sample (i.e., in any one run) are useful indications of the expected sampling error in the point estimates.

Table 1: Monte Carlo Results for Basic Specification

	Alternative-specific constant	Fixed coefficient	Random coefficient: mean	Random coefficient: standard dev.	Scale
True value	1	1	1	0.5	4
Point estimates: Mean	0.9940	0.9988	0.9984	0.4843	4.3736
Standard deviation	0.0753	0.0961	0.0586	0.1434	2.5449
Standard errors: Mean	0.0769	0.0950	0.0639	0.1383	2.7226
Standard deviation	0.0038	0.0068	0.0108	0.0314	4.7470

An important issue is whether, or the extent to which, the sp-off-rp data provide better estimates than the rp data alone. Table 2 summarizes the results of estimation on the rp data alone (i.e., the agent's choice between alternatives 1 and 2 based on U_{1n} and U_{2n}) without the sp-off-rp data. The mean estimates are close to the true values, though for two of the parameters (the fixed coefficient and the mean of the random coefficient) the hypothesis that the mean equals the true value can be rejected at the 95% level. The standard deviations of the estimates are larger using only the rp data than when using the combined rp and sp-off-rp data. The sp-off-rp exercise changes the value of x, which has the random coefficient, and, not surprisingly, the greatest effect is observed in the estimated parameters of the random coefficient. The use of the sp-off-rp data (Table 1) reduces the standard deviation of the estimates obtained from the use of rp data (Table 2) by over half, from 0.1361 to 0.0586 for the mean of the random coefficient and from 0.3821 to 0.1434 for the estimated standard deviation. To put this improvement in perspective, using the sp-off-rp data is equivalent to more than quadrupling sample size with rp data alone[5] (since a four-fold increase in sample size reduces asymptotic standard errors by two.) Interestingly, the estimates of the fixed coefficient and the alternative-specific constant are also improved by the sp-off-rp data, even though the sp-off-rp exercise only changed the variable

[5] Given the cost of sampling, the result suggests significant savings of approximately 75 percent.

with the random coefficient. The standard deviations of these parameters drop by more than 20% when using the sp-off-rp data. The sp-off-rp data allow more precise estimation of these parameters because an agent's response to a change in one variable (i.e., the one changed in the sp-off-rp exercise) depends on the difference in the utility between alternatives prior to the change and, therefore, reveals information about all utility parameters.

Table 2: Monte Carlo Results for Basic Specification, using Revealed-Preference Data Only

	Alternative-specific constant	Fixed coefficient	Random coefficient: mean	Random coefficient: standard dev.
True value	1	1	1	0.5
Point estimates: Mean	1.0098	1.0278	1.0502	0.5473
Standard deviation	0.0988	0.1208	0.1361	0.3821
Standard errors: Mean	0.0988	0.1207	0.1496	0.5691
Standard deviation	0.0139	0.0147	0.0349	0.1666

A complication inherent in using sp-off-rp data is the need to model endogeneity. To see the effect of ignoring the endogeneity inherent in the sp-off-rp data, estimation was performed as if the sp-off-rp data were the same as standard sp data. This procedure is denoted estimation on "sp" data, in quotes. The agent's choice under the changed value of x was modelled as a standard mixed logit, with mixing over the distribution of the random coefficient but without including the conditional error terms. The rp and "sp" data were combined for joint estimation, and a separate scale was allowed for the "sp" choice, as is customary when combining rp and sp data. The results are summarized in Table 3. The mean estimates are all significantly different from the true values. The differences are most prominent in the standard deviation of the random coefficient (whose mean estimate is more than twice the true value) and the scale parameter (whose mean estimate is less than a quarter of the true value.) It is not clear why the error is more concentrated in these parameters than the others, and we have not investigated whether the pattern arises under other specifications. The result simply shows that estimation on sp-off-rp data as if they were standard sp data can cause substantial estimation error, and points to the need to model the endogeneity.

Table 3: Monte Carlo Results for Basic Specification, ignoring endogeneity

	Alternative-specific constant	Fixed coefficient	Random coefficient: mean	Random coefficient: standard dev.	Scale
True value	1	1	1	0.5	4
Point estimates: Mean	1.1288	1.1513	0.9555	1.2445	0.6379
Standard deviation	0.0942	0.1236	0.1142	0.2026	0.0893
Standard errors: Mean	0.0934	0.1131	0.1184	0.1982	0.0971
Standard deviation	0.0061	0.0084	0.0143	0.0206	0.0127

Table 4 summarizes results with estimation on 250 observations instead of 1000. The top part of the panel gives results for estimation on the rp and sp-off-rp data, while the bottom part has results for estimation on the rp data alone. Since sample size is reduced by four, the standard deviations of the estimates and the mean standard errors are expected to double, provided that the smaller sample size is still sufficiently large for the asymptotic distributions to be approximately accurate. As can be seen in the top part of the table, the mean estimates are close to the true values, with no significant differences even with the smaller sample size. The standard deviations are similar to the mean standard errors, and both are about twice as large as their values in Table 1 with 1000 observations. These results imply that the asymptotical distribution still seems to serve as a good approximation with as few as 250 observations.

The results on the rp data alone indicate that the sp-off-rp data are even more useful for small samples than for the larger sample. First, the standard deviation of the estimates based on rp data alone increase by considerably more than twice when reducing sample size from 1000 to 250 (next to last row of Table 4 compared with middle row of Table 2.) Second, even with these larger standard deviations, the mean estimates are significantly different from their true values for three out of the four parameters (third from last row of Table 4.) These two results indicate that the sample size is too small for the asymptotic properties to be exhibited when estimation is performed on the rp data alone. The inclusion of the sp-off-rp data reduces by a factor of over three the standard deviations of the estimates of the parameters of the random coefficients (top part of Table 4 compared with bottom part). This improvement is greater than the two fold improvement that was obtained with a sample of 1000, discussed above.

Table 4: Monte Carlo Results for Basic Specification, 250 observations

	Alternative-specific constant	Fixed coefficient	Random coefficient: mean	Random coefficient: standard dev.	Scale
True value	1	1	1	0.5	4
RP and SP-off-RP Data					
Mean estimates	1.0073	1.0054	1.0220	0.4846	4.6784
Std dev. estimates	0.1399	0.1741	0.1252	0.2784	6.4835
Mean Ses	0.1583	0.1911	0.1250	0.2812	5.8497
RP data alone					
Mean estimates	1.0813	1.0800	1.1854	0.8262	
Std dev. estimates	0.2460	0.3013	0.4646	0.9943	
Mean Ses	0.2185	0.2638	0.3523	1.0448	

We next examine various aspects of the specification to assess the impact of each element on the efficiency of the estimator. In particular, we make each of the following changes in specification:

a. Reduce the range of the explanatory variables to be uniform between 2.5 and 3.5 instead of 2 and 4, such that the difference between alternatives ranges from -1 to +1 instead of -2 to +2.

b. Reduce the level of the explanatory variables to be uniform between 1 and 3 instead of 2 and 4. The difference between alternatives still ranges from -2 to +2. The reduction in level changes the magnitude of the reduction in x for the chosen alternative in the sp-off-rp data. (Since x is reduced by a proportion of its value, the reduction is smaller in magnitude when the level of x is smaller.)

c. Reduce the range of reductions in x, such that the proportion reduction r_n is uniformly distributed between 0.25 and 0.75 instead of 0 to 1.

d. Reduced the scale from 4 to 2, thereby doubling the standard deviation of the error associated the sp-off-rp choice.

e. Reduce the scale even further to 0.5, thereby increasing the standard deviation of the processing error by a factor of eight relative to the original specification and by a factor of four relative to the specification in (d).

Each of these changes is designed to decrease the efficiency of the estimator by decreasing either the variation in the data (specifications a-c) or the precision of the agents' responses to the sp-off-rp question (specifications d and e.) Table 5 summarizes the results. The mean estimates are given in the top part of the table and the standard deviations in the bottom. For comparison, the first row of each part gives results for the original specification (i.e., repeats the information from Table 1.)

The means are close to the true values in all specifications. Using a t-test at the 95% confidence level, the hypothesis that the mean is equal to the true value is rejected in only four instances, whose means are given in bold in the table. Since there are a total of 30 such tests, the expected number of rejections when the hypothesis is true is 1.5, and the probability of obtaining 4 or more rejections is 0.06. The hypothesis that all the means are equal to their true values can, therefore, be rejected at the 95% level but not that the 97% level. In any case, the differences are small and the significant ones are not concentrated in any one specification.

Table 5: Monte Carlo Results for Variations on Basic Specification

	Alternative-specific constant	Fixed coefficient	Random coefficient: mean	Random coefficient: standard dev.	Scale
Mean estimates					
Base specification	0.9940	0.9988	0.9984	0.4843	4.3736
a. X and Z uniform 2.5-3.5	0.9977	1.0071	0.9914	**0.4509**	4.7696
b. X and Z uniform 1-3	0.9948	0.9877	**0.9853**	0.4740	4.8225
c. R_n uniform .25-.75	1.0139	**1.0272**	1.0058	0.4767	4.3213
d. Scale =2	0.9915	**0.9805**	0.9912	0.4672	2.0782
e. Scale = 0.5	0.9902	0.9813	0.9869	0.4455	0.4966
Standard deviations					
Base specification	0.0753	0.0961	0.0586	0.1434	2.5449
a. X and Z uniform 2.5-3.5	0.0712	0.1480	0.0713	0.1849	4.1215
b. X and Z uniform 1-3	0.0739	0.0880	0.0675	0.1854	4.7024
c. R_n uniform .25-.75	0.0719	0.1058	0.0725	0.1609	5.3613
d. Scale =2	0.0770	0.0912	0.0703	0.1821	0.6913
e. Scale =0.5	0.0851	0.1050	0.0875	0.3702	0.0910

Specification (a) reduces the range of the explanatory variables relative to the base specification. As expected, the standard error of the parameters associated with both variables, as well as the scale parameter, rise relative to those in the base specification. In specification (b), the level of x and z for each alternative is decreases by 1. This change does not affect the difference in variables between alternatives in the rp choice, since the reduction is applied to each alternative. However, in the sp-off-rp data, the value of x for the chosen alternative is reduced by a proportion, while the value of x for the nonchosen alternative is not changed. The effect of the new specification, therefore, is to reduce the range of x in the sp-off-rp question. As expected, the standard deviations of the parameters for the coefficient of x and the scale of the sp-off-rp error rise. Specification (c) also decreases the range of x in the sp-off-rp data, by decreasing the range of the proportion by which x for the chosen alternative is reduced. As with specification (b), the standard deviations of the parameters of the coefficient of x and the scale parameter rise.

Specification (d) and (e) increase the standard deviation of the additional error that enters the sp-off-rp choices, which, intuitively, makes these choices more "noisy" and, hence, less useful for estimation of the true behavioral parameters. The scale is estimated fairly precisely in each case: the mean estimate is 2.0782 when scale is 2.0, and 0.4966 when scale is 0.50. The standard deviations rise, as expected, but far less than the increase in the standard deviation of the error. For example, the standard deviation of the estimates of the standard deviation of the coefficient of x (which is the parameter that is most affected by the change in scale) rises from 0.14 to 0.18 when the standard deviation of the error doubles, and rises from 0.14 to 0.37 when the standard deviation of the error rises by a factor of eight.

An important issue is whether, or the conditions under which, sp-off-rp data provide more information for estimation than standard sp experiments. To address this issue, simulations were performed with each agent presented with a standard sp experiment rather than an sp-off-rp experiment. For the first comparison, each agent is given a choice between two alternatives that differ in x and the identity of the alternative (which determines the alternative-specific constant.) This set-up, with only x varying, corresponds to the sp-off-rp choice in which the value of x was changed. Specifications with both x and z varying are considered below. Utility in the sp choice is assumed to take the same form as in the rp choice, with each agent using the same parameters as in their rp choice, except that the standard deviation of the error in the sp choice differs from that in the rp choice by a factor $(1/\lambda)$. The model was estimated on the combined rp and sp data, with a separate scale λ for the sp data and all other parameters being the same. Estimation of a separate scale for rp and sp data when combining the two is standard practice; see, e.g., Ben-Akiva and Morikawa (1990), Hensher and Bradley (1993), Louviere *et al*. (2000), and Train (2003, section 7.2). It is also standard practice to estimate separate alternative-specific constants on the rp and sp data. We instead use the same constant for both types of data (both in simulation of the choices and in estimation), which increases the efficiency of the sp estimator in our analysis. The standard deviations of the parameter estimates on the rp/sp data in our analysis are, therefore, smaller than would be expected under standard practice.

The results are summarized in Table 6, which, for comparison, also contains results for estimation using the sp-off-rp data (repeated from previous tables.) Estimation is performed with true scale set at 4, 2 and 0.5, with smaller scale indicating greater processing error in the sp choices (i.e., larger standard deviation of the unobserved portion of utility in the sp choices). The scale parameter for the sp data is not exactly comparable to the scale parameter for the sp-off-rp data. For the sp data, the scale reflects the standard deviation of the unobserved portion of utility in the sp choice relative to that in the rp choice. For the sp-off-rp data, the scale reflects the standard deviation of the *extra* error that is added to the unobserved portion of utility in the rp setting. The same value of the scale parameter, therefore, implies larger total error in the sp-off-rp utility than in the sp utility. This difference in the meaning of the scale parameter implies that the comparisons in Table 6 are biased in favor of the sp data over the sp-off-rp data, since the set-up gives a larger error for the sp-off-rp data than the sp data.

Table 6: Monte Carlo Results for SP Data and SP-off-RP Data

	Alternative-specific constant	Fixed coefficient	Random coefficient: mean	Random coefficient: standard dev.	Scale
Mean estimates					
Scale = 4					
SP-off-RP	0.9940	0.9988	0.9984	0.4843	4.3736
SP	1.0058	1.0049	1.0005	0.4920	**4.1270**
Scale = 2					
SP-off-RP	0.9915	**0.9805**	0.9912	0.4672	2.0782
SP	1.0061	1.0050	1.0001	0.4691	2.0130
Scale = 0.5					
SP-off-RP	0.9902	0.9813	0.9869	0.4455	0.4966
SP	1.0081	1.0107	1.0082	**0.5809**	**0.4850**
x and z vary					
SP-off-RP	0.9965	1.0039	0.9997	0.4884	**3.8364**
SP	1.0058	1.0066	0.9932	0.4895	4.0210
Standard deviations					
Scale = 4					
SP-off-RP	0.0753	0.0961	0.0586	0.1434	2.5449
SP	0.0819	0.1063	0.0974	0.1305	0.5883
Scale = 2					
SP-off-RP	0.0770	0.0912	0.0703	0.1821	0.6913
SP	0.0848	0.1073	0.1002	0.1950	0.2079
Scale = 0.5					
SP-off-RP	0.0851	0.1050	0.0875	0.3702	0.0910
SP	0.0939	0.1138	0.1260	0.3473	0.0685
x and z vary					
SP-off-RP	0.0735	0.0442	0.0486	0.0743	0.5672
SP	0.0691	0.0706	0.0756	0.0677	0.4038

The mean estimates based on combined rp and sp data are similar to the true values, with the hypothesis of equality to the true value being rejected only three times in the 20 tests (shown in bold.) In this regard, the sp data perform about the same as the sp-off-rp data, which obtained two rejections out of the 20. For a given level of the scale parameter, the standard deviations of the estimates using sp data are similar to those using sp-off-rp data, with some being smaller and some larger. As the scale drops (i.e., as the "processing" error becomes greater), the standard deviations of the estimates rise under both approaches. These two results combined imply that the procedure that has the lower processing error can be expected to provide more precise estimates. One of the motivations for the use of sp-off-rp questions instead of sp experiments is that, by asking questions in relation to the respondent's a real-world choice, the respondent is more able to meaningfully assess the hypothetical situation. If this conjecture is true, or, more precisely, if the processing error in sp-off-rp choices is, indeed, less than in sp experiments, then these simulation results indicate that greater estimation efficiency is obtained with sp-off-rp data than with sp data.

Table 6 contains one last comparison. In the specifications considered so far, only x was varied in the sp-off-rp and sp data. It is, of course, customary to include a series of sp-off-rp or sp tasks with each relevant variable varying. We next consider, therefore, sp-off-rp questions about both z and x, and sp experiments that contain both variables. The specification is the same as the base specification, with scale parameter of 4. For the sp-off-rp data, each agent is asked two questions: one question (the same as in earlier specifications) about how they would respond if x for their chosen alternative were reduced by a certain proportion, and a second question that is similar but for a reduction in z for their chosen alternative. The outcome consists of the agent's choice between the original rp alternatives, their choice when x for their chosen rp alternative is reduced, and their choice when z for their chosen rp alternative is reduced. Sp data are specified analogously. Two sp experiments are administered for each agent, with x and z varying over alternatives and experiments. The outcome consists of the agent's choice between the rp alternatives and their choices in the two sp experiments.

The last rows in both parts of Table 6 summarize the results for these specifications. The standard deviations are considerably lower than with only x varying. For example, the standard deviation of the fixed coefficient of z drops from 0.0961 when asking an sp-off-rp question about x only to 0.0442 when asking questions about both x and z. Similarly, for sp experiments, the standard deviation drops from 0.1063 using one experiment with x varying to 0.0706 using two experiments with both x and z varying. The standard deviations are about the same for the two methods when both x and z are varied, with the sp-off-rp data obtaining a lower standard deviation for some parameters (viz., the fixed coefficient and the mean of the random coefficient) and the sp data obtaining smaller standard deviations for the other parameters (the intercept, scale, and standard deviation of the random coefficient.) These results confirm the earlier statement based on one sp-off-rp and sp task that, when the processing error for the two types of data are the same, sp-off-rp questions and sp experiments provide about the same level of estimation efficiency. The researcher's decision of which method to use depends largely, therefore, on which method the researcher expects will induce less processing error by the respondents.

3. Conclusions

Sp-off-rp data are generated by changing the attributes of alternatives in an rp setting on the basis of the agent's choice in that setting. The primary advantages of such data are that: 1) as with any form of sp tasks, the data can contain substantially more variation in the attributes underlying the choice than is commonly observed in rp data; and 2) the sp-off-rp data are constructed from the revealed choice made by the agent and, as such, overcome the common criticism of sp data i.e., the lack of realism. Yet, since the sp-off-rp data are endogenous, estimation is more complicated. The Monte Carlo results presented in this paper suggest that the added complication is repaid in potentially substantial gains in efficiency. In our base specification, models estimated with sp-off-rp data obtained approximately the same level of efficiency as models estimated with rp data, but with only about ¼ of the observations. This approach, therefore, can provide substantial savings in sampling costs.

Responses to the sp-off-rp questions may differ for a variety of unobserved factors that are unrelated to the rp error; e.g., the respondent may not be attentive to the task or may tend to answer randomly. Such issues also arise in standard sp experiments. The estimator for sp-off-rp designs explicitly allows for quixotic responses. As expected, the efficiency gain from sp-off-rp data rises when this error variance declines. Importantly, our Monte Carlo results indicate that sp-off-rp data provide greater efficiency than standard sp data if, as expected, the variance of this response error is lower in sp-off-rp data than in sp data.

Finally, the Monte Carlo experiments suggest that it is critically important to model the endogeneity in sp-off-rp data. Indeed, if one uses sp-off-data to estimate the parameters of a choice model, but ignores the endogenous construction of the data, significant bias can be introduced. Thus, for a given sample size, there are potential efficiency gains from an sp-off-rp design, but these gains can only be attained when the estimation procedure appropriately reflects the experimental design.

References

Anas, A. and C. Feng, 1988, "Invariance of expected utilities in logit models," *Economic Letters*, Vol. 27, pp. 57-78.

Ben-Akiva, M. and T. Morikawa, 1990, "Estimation of switching models from revealed preference and stated intentions," *Transportation Research A*, Vol. 24, pp. 485-495.

Bhat, C. and S. Castelar, 2001, "A unitfied mixed logit framework for modeling revealed and stated preferences: Formulation and application to congestion pricing in the San Francisco bay area," *Transportation Research B*, Vol. 36, pp. 577-669.

Bradley, M. and A. Daly, 2000, "New analysis issues in stated preference research," in J. de D. Ortuzar, ed., *Stated Preference Modeling Techniques*, PTRC Education and Research Services, London.

Brownstone, D., D. Bunch and K. Train, 2000, "Joint mixed logit models of stated and revealed preferences for alternative-fueled vehicles," *Transportation Research B*, Vol. 34, pp. 315-338.

Caussade, S., J. de D. Ortuzar, L. Rizzi and D. Hensher, 2005, “Assessing the influence of design dimensions on stated choice experiment estimates,” *Transportation Research B*, Vol. 39, pp. 621-640.

Fowkes, A. and N. Shinghal, 2002, The Leeds adaptive stated preference methodology,” in R. Daniels, ed., *Freight Transportation Demand and Stated Preference Experiments*, FrancoAngeli, Milan.

Greene, W., D. Hensher and J. Rose, 2006, “Accounting for heterogeneity in the variance of unobserved effects in mixed logit models,” *Transportation Research B*, Vol. 40, pp. 75-92.

Hensher, D., 2004, “Accounting for stated choice design dimensionality in willingness to pay for travel time savings,” *Transportation Research B*, Vol. 38, pp. 425-446.

Hensher, D. and M. Bradley, 1993, “Using stated response data to enrich revealed preference discrete choice models,” *Marketing Letters*, Vol. 4, pp. 39-152.

Hensher, D., J. Louviere and J. Swait, 1999, “Combining sources of preference data,” *Journal of Econometrics*, Vol. 89, pp. 197-221.

Hensher, D. and J. Rose, 2007, “Development of commuter and noncommuter mode choice models for the assessment of new public transport infrastructure projects: A case study,” *Transportation Research A*, Vol. 41, pp. 428-443.

Louviere, J., D. Hensher and J. Swait, 2000, *Stated Preference Methods: Analysis and Applications*, Cambridge University Press, New York.

Rose, J., M. Bliemer, D. Hensher and A Collins, forthcoming, “Designing efficient stated choice experiments in presence of reference alternatives,” *Transportation Research B*.

Train, K., 2003, *Discrete Choice Methods with Simulation*, Cambridge University Press, New York.

Train, K. and W. Wilson, 2007, "Estimation on stated preference experiments constructed from revealed-preference choices," forthcoming, *Transprotation Research B*.

navigation · economics · technologies

The NETS research program is developing a series of practical tools and techniques that can be used by Corps navigation planners across the country to develop consistent, accurate, useful and comparable information regarding the likely impact of proposed changes to navigation infrastructure or systems.

The centerpiece of these efforts will be a suite of simulation models. This suite will include:

- A model for forecasting international and domestic traffic flows and how they may be affected by project improvements.
- A regional traffic routing model that will identify the annual quantities of commodities coming from various origin points and the routes used to satisfy forecasted demand at each destination.
- A microscopic event model that will generate routes for individual shipments from commodity origin to destination in order to evaluate non-structural and reliability measures.

As these models and other tools are finalized they will be available on the NETS web site:

http://www.corpsnets.us/toolbox.cfm

The NETS bookshelf contains the NETS body of knowledge in the form of final reports, models, and policy guidance. Documents are posted as they become available and can be accessed here:

http://www.corpsnets.us/bookshelf.cfm

NETS

navigation · economics · technologies