A CONTROL-FUNCTION APPROACH TO CORRECT FOR ENDOGENEITY IN DISCRETE CHOICE MODELS ESTIMATED ON SP-OFF-RP DATA

C. Angelo GUEVARA

Departamento de Ingeniería Civil,
Universidad de Chile,
crguevar@ing.uchile.cl,
+56229784380,
Blanco Encalada 2002, Santiago, Chile

Stephane HESS

Institute for Transport Studies & Choice Modelling Centre,
University of Leeds,
s.hess@its.leeds.ac.uk,
+44 113 343 6611, 36-40
University Road, LS2 9JT, Leeds

4,534 words + 4 tables and figures.

total 5,534

August 1, 2017

ABSTRACT

It is common practice to build Stated Preference (SP) attributes and alternatives from observed Revealed Preference (RP) choices. While many surveys pivot all alternatives around an observed choice, others use more adaptive approaches in which changes are made depending on what alternative was chosen in the RP setting. For example, in SP-off-RP data, the RP chosen alternative is worsened and other alternatives are improved to induce a choice change. This facilitates the creating of meaningful trade-offs or tipping points but introduces endogeneity. This source of endogeneity was largely ignored until Train and Wilson (T&W) proposed a full information maximum likelihood (FIML) solution that can be implemented with simulation. In this article we propose a limited information maximum likelihood (LIML) approach to address the SP-off-RP problem using a method that does not need simulation, can be applied with standard software and uses data that is already available for the stated problem. The proposed method can be seen as an application of the Control-Function (CF) method to correct for endogeneity in discrete choice models, using the RP attributes as instrumental variables. We discuss the theoretical and practical advantages and disadvantages of the CF and T&W methods and illustrate them using Monte Carlo and real data. We also suggest that the T&W results are driven by accounting for correlation across SP choices.

Keywords: Stated-preference, Revealed Preference, Endogeneity.

1. INTRODUCTION

Stated Preferences (SP) methods are an important tool because they allow the presentation of non-existing alternatives, control the variability of the data and are inexpensive. To maximize the realism and the information gathered from the experiment the choice set presented is often customized for a respondent based on their revealed preferences.

A very popular approach consists of using pivoting around a revealed preference (RP) choice (1), where a number of hypothetical alternatives are presented that involve increases and decreases around the real world choice for all alternatives. While this approach can avoid endogeneity bias as the changes are made to all alternatives irrespective of the choice observed in the RP data, it does not guarantee that the trade-offs are tailored to create tipping points that will allow an analyst to understand at what point the decision maker will move away from the RP choice.

An alternative approach consists of using adaptive design approaches where the original choice set is reproduced in the SP setting but where the attribute levels are changed as a function of the RP choice. In particular, the variables of the alternative chosen in the RP setting are made “worse”, while those of the remaining alternatives are made “better”. Examples of such approaches include the adaptive SP approaches of (2) and more recently the SP-off-RP approach (3,4), where, in SP choices, the RP chosen alternative is worsened and other alternatives are improved to induce a choice change.

A crucial problem arises as this approach creates an endogeneity risk which, if uncontrolled, will lead to inconsistent estimates. Since the choice-set presented to the individual depends on the RP choices, and therefore on that individual’s preferences, an estimation method that neglects this endogenous setting would result in inconsistency. Correction approaches exist but they are difficult to apply. The work of (3), hereafter T&W, offered a Full Information Maximum Likelihood (FIML) solution for the case of SP-off-RP data that can be implemented with simulation. This seminal work has been widely cited but rarely applied, arguably because of its various practical difficulties. As a result, SP-off-RP surveys have also failed to be widely applied.

In this article we propose a Limited Information Maximum Likelihood (LIML) approach to address the SP-off-RP problem using a method that does not need simulation, can be applied with standard software and use data that is already available for the stated problem.

The remainder of this paper is organised as follows. The next section presents the SP-off-RP design approach and the T&W correction for endogeneity. We follow this in Section 3 by our proposed CF method. Empirical illustrations follow in Section 4 using simulated data and Section 5 using real data. Finally, we present some conclusions and directions for future work.

2. SP-OFF-RP DATA AND TRAIN & WILSON CORRECTION FOR ENDOGENEITY

The SP-off-RP problem is a two stages process. The first stage corresponds to an RP choice that occurs in the real environment in which individual *n* chooses an alternative i_n^{RP} among the choice set C_n^{RP} by maximizing the utility $U_{in}^{RP} = V_{in}^{RP} + \varepsilon_{in}^{RP} = \sum_k \beta_k x_{ink}^{RP} + \varepsilon_{in}^{RP}$ perceived from each alternative.

This utility depends on a systematic part V_{in}^{RP} and random part ε_{in}^{RP}. The systematic part is formed a as linear combination of attributes x_{ink}^{RP} with, coefficients β_k. The random part is assumed to be independently and identically distributed Extreme Value I, with location zero and scale μ^{RP}. Under these assumptions, the probability that individual *n* would choose alternative *i* in the RP experiment will be $P_n^{RP}(i)$, as shown in Eq. [1].

$$\begin{aligned}
&U_{in}^{RP} = V_{in}^{RP} + \varepsilon_{in}^{RP} \qquad \varepsilon_{in}^{RP}\ iid\ \mathrm{EV}\left(0, \mu^{RP}\right) \\
&y_{in}^{RP} = 1\left[U_{in}^{RP} \geq U_{jn}^{RP} \forall j \in C_n^{RP}\right] \\
&i_n^{RP} : \text{RP Choice of individual } n \\
&P_n^{RP}(i) = \frac{e^{\mu^{RP} V^{RP}(x_{in}, \beta)}}{\sum_{j \in C_n^{RP}} e^{\mu^{RP} V^{RP}(x_{jn}, \beta)}} \qquad \mu^{RP} \equiv 1
\end{aligned} \qquad [1]$$

In the second stage of the SP-off-RP choice process, the researcher builds an SP experiment which considers a choice set D_n^{SP} that includes i_n^{RP}, the alternative chosen in the RP experiment, and a set (possibly a subset) of other alternatives from C_n^{RP}. The attributes of the resulting set of SP alternatives are built as variations of their values in the RP experiment in a way that tries to induce a choice change. For example, the attribute x_{jnk}^{SP} of alternative *j* and individual *n* in the SP experiment could calculated be as follows

$$x_{jnk}^{SP} = \gamma_{jnk} x_{jnk}^{RP}, \qquad [2]$$

where γ_{jnk} is an endogenous scalar defined by the researcher, such that if $\beta_k < 0$, then

$$\gamma_{jnk} \begin{cases} > 1 \text{ if } j = i_{i_n^{RP} n} \\ < 1 \text{ if } j \neq i_{i_n^{RP} n} \end{cases}. \qquad [3]$$

for a given n. This implies that for undesirable attributes (i.e. $\beta_k < 0$), the level of the attribute is increased for that alternative which was chosen in the RP setting while it is decreased for all other alternatives. For desirable attributes, the opposite applies. The net outcome of this is that the alternative chosen in the RP setting becomes less attractive, while those that were not chosen become more attractive.

Under this framework, it is plausible to think that the choice-maker will transfer some part of the unobserved attributes of each alternative from the RP to the SP experiment, for the same reason that we assume realism increases by building SP experiments that closely related to RP. This transferability can be accommodated by considering that some part of ε_{jn}^{RP} will appear also in the utility of the alternative of the SP choice model. Under that framework, the SP choice model could be written as follows

$$\begin{aligned} &U_{jn}^{SP} = V\left(x_{jn}^{SP}, \beta\right) + \rho\varepsilon_{jn}^{RP} + \varepsilon_{jn}^{SP} \qquad \varepsilon_{jn}^{RP}, \varepsilon_{jn}^{SP} \; iid \; \mathrm{EV}\mu^{\mathrm{RP}}, \mu^{\mathrm{SP}} \\ &y_{jn}^{SP} = 1\left[U_{jn}^{SP} \geq U_{\mathrm{ln}}^{SP} \forall l \in C_n\right] \qquad , \\ &j_n^{SP} : \text{SP Choice of individual } n \end{aligned} \quad [4]$$

where ρ corresponds to the fraction of the RP error that was transferred to the SP experiment. In more recent work, (5) remark that (3) implicitly assume that $\rho = 1$, something that is not necessarily true.

Whenever $\rho \neq 1$, endogeneity problems will arise because ε_{jn}^{RP} is correlated with x_{jn}^{SP} by Eq. [2] and [3], as γ_{jn} depends on the RP choice and, thus, on ε_{jn}^{RP}. Therefore, consistency in the estimation of the model coefficients will not be achieved if an analyst ignores this transfer of error, i.e. estimates a model either on just the SP data or jointly on the SP and RP data, without creating a link between the two.

The work of (3) put forward a solution to this problem where the probability of alternative *j* in the SP setting, conditional on *i* being chosen in the RP setting is given by a mixed logit model as:

$$P_{n,k|i} = \int \frac{e^{\alpha\beta' x_k + \rho\alpha\epsilon_k}}{\sum_j e^{\alpha\beta' x_j + \rho\alpha\epsilon_j}} f\left(\epsilon \mid \beta' x_i + \epsilon_i > \beta' x_j + \epsilon_j, \forall j \neq i \right) d\epsilon, \quad [5]$$

where the integration is over the conditional on density of ϵ, conditional on *i* being chosen in the RP setting. Joint estimation on the RP and *T* separate SP choices then maximises:

$$P_n = \int \prod_{t=1}^{T} \left(\frac{e^{\alpha\beta' x_{k_t} + \rho\alpha\epsilon_{k_t}}}{\sum_j e^{\alpha\beta' x_j + \rho\alpha\epsilon_j}} \right) f\left(\epsilon \mid \beta' x_i + \epsilon_i > \beta' x_j + \epsilon_j, \forall j \neq i \right) d\epsilon \, \frac{e^{\beta' x_i}}{\sum_j e^{\beta' x_j}}, \quad [6]$$

where alternative k_t is chosen in task *t*. A further level of flexibility is added by not setting ρ to 1. The estimate for α is then a scale parameter for SP data while τ gives an indication of how much of the RP error is transferred to the SP setting.

The actual way to produce draws from ϵ is not computationally difficult albeit tedious to implement given that they come from an extreme value distribution with a shifted mean (by the negative logarithm of the probability of the RP choice) for the alternative chosen in the RP setting and a truncated extreme value distribution for those alternatives not chosen in the RP setting (truncated by the differences in the deterministic utilities for *i* and *j* plus the error term for alternative *i*). Even if the model the analyst wishes to estimate is MNL, simulation based estimation is required with this method.

3. PROPOSED CONTROL FUNCTION CORRECTION FOR THE SP OFF RP PROBLEM

In this article, we propose a method to correct for the endogeneity problem in the SP-off-RP problem taking advantage of the fact that RP attributes are suitable instrumental variables for the SP attributes, what allows applying the control-function (CF) approach. The method we propose has the advantage that is much easier to apply but with a potential cost in efficiency. We think that this method may be an important step toward transforming this critical correction for endogeneity into a mainstream method and lead to wider use of SP-off-RP surveys.

The CF method was originally proposed by (6), with later contributions by (7,8,9,10). The application of CF for the correction of endogeneity requires an instrumental variable for each endogenous variable of the model. This instrumental variable has to be, at the same time, sufficiently correlated with the respective endogenous variable and independent of the error term of the model.

Because of Eq. [2], the endogenous variables in this case are any x_{jn}^{SP} build as a variation of the respective x_{jn}^{RP}, because γ_{jn} depends on the RP choice and then on the transferred ε_{jn}^{RP}. Also because of Eq. [2], x_{jn}^{RP} makes proper instrument for x_{jn}^{SP} because they are correlated among them, and because x_{jn}^{RP} is independent of ε_{jn}^{RP}.

Therefore, we can estimate the SP model alone if we correct for the endogeneity using the two steps control-function (CF) approach, which is much easier to apply than the current state of the art and has the advantage that it can be applied using standard software. On the negative side, the proposed method may be less efficient as it does not account for correlation across the different SP choices and the standard errors cannot be calculated directly from the inverse of the Fisher Information Matrix. However, it can be said that the potential lower efficiency comes also at the benefit of being more robust to the distributional assumptions and that the additional difficulty in the estimation of the standard errors has been already circumvented analytically (11), or can be easily handled by non-parametric methods.

To explain the method, consider a binary mode choice model between car and bus and with only two x_{jn}^{RP} attributes: travel time and cost. Since people would like to travel for a shorter time and paying a smaller amount of money, $\beta_t < 0$ and $\beta_c < 0$. Then, if, for example, individual *n* chooses the bus in the RP experiment, the attributes of SP experiment will be built to try to induce a choice change, worsening the bus and improving the car. This could be achieved if, for example,

$$\begin{aligned} Time_{Bus\_n}^{SP} &= 1.1 Time_{Bus\_n}^{RP} \\ Cost_{Bus\_n}^{SP} &= 1.3 Cost_{Bus\_n}^{RP} \\ Time_{Car\_n}^{SP} &= 0.8 Time_{Car\_n}^{RP} \\ Cost_{Car\_n}^{SP} &= 0.7 Cost_{Car\_n}^{RP} \end{aligned}, \qquad [7]$$

where the scalars γ_{jnk} could have been fully predefined or randomly drawn from a pre-defined set of scalars that are larger or smaller than one, as it corresponds.

Then, the CF correction is applied in two stages. In the first stage, each endogenous x_{jn}^{SP} has to be regressed on their respective instrument and the controls, to retrieve a residual that will capture the part of x_{jnk}^{SP} that was correlated with the error of the SP model. Because of Eq. [7] the instrument for each x_{jnk}^{SP} is its respective x_{jnk}^{RP}. Because in this example we have more than one endogenous variable, the regression has to be made not only on the respective instrument, but also on the instruments of the other endogenous variables, which act as controls. For the stated example, this means estimating the two regressions shown in Eq. [8] by ordinary least squares, stacking the information from all available alternatives.

$$\begin{aligned} Time_{j\_n}^{SP} &= \alpha_0 + \alpha_t Time_{j\_n}^{RP} + \alpha_t Cost_{j\_n}^{RP} + \delta_{jn\_time} \\ Cost_{j\_n}^{SP} &= \alpha_0 + \alpha_t Time_{j\_n}^{RP} + \alpha_t Cost_{j\_n}^{RP} + \delta_{jn\_\cos t} \end{aligned} . \quad [8]$$

Then, in the second stage, the residuals of these regressions $\hat{\delta}_{jn\_time}, \hat{\delta}_{jn\_\cos t}$ are added to the SP choice model to control for endogeneity, such that the systematic part of the utility takes the following form:

$$V_{jn}^{SP} = ASC_j + \beta_{time}\mathrm{time}_{jn}^{SP} + \beta_{\cos t}\mathrm{cost}_{jn}^{SP} + \hat{\theta}_{time}\hat{\delta}_{jn\_time} + \hat{\theta}_{\cos t}\hat{\delta}_{jn\_\cos t} \quad [9]$$

The statistical significance of $\hat{\theta}_{time}, \hat{\theta}_{\cos t}$ can be used as a test for the presence of endogeneity and, under the setting described, the proposed correction will allow the consistent estimation of the ratio of the model coefficients $ASC_j, \beta_{time}, \beta_{\cos t}$ (9).

As discussed in (9), the two stage CF approach has the advantage of being applicable with canned software, but at the cost of reducing efficiency, compared to FIML methods, such as the one proposed by (3) for the problem under analysis. In addition, standard errors in the CF method have to be calculated using non-parametric methods such as the bootstrap (8) or the delta-method (11).

4. MONTE CARLO EXPERIMENT CONTRASTING THE CF AND T&W APPROACHES

In this section we report a Monte Carlo Experiment to illustrate the application of the proposed CF method contrasted with the current state of the art for addressing endogeneity in the SP-off-RP problem.

The data generation process used to build the RP experiment in this Monte Carlo was a trinary mode choice model of 1,000 observations that depended solely on travel time and travel cost with alternative specific constants set to zero and using the population coefficients shown in Eq. [10].

$$U_{in}^{RP} = -1Time_{j\_n}^{RP} - 0.5Cost_{j\_n}^{RP} + \varepsilon_{jn}^{RP} . \quad [10]$$

$Time_{j\_n}^{RP}$ and $Cost_{j\_n}^{RP}$ values were drawn from a random uniform between 1 and 3. The error ε_{jn}^{RP} was drawn from an Exteme Value I (Gumbel) with location zero and scale $\mu^{RP} = 1$, such that, if the individual selects the alternative with the largest utility, the resulting model becomes a Logit.

An SP experiment was built for each individual by shifting the attributes of the RP experiment. The multiplier shift was randomly drawn from a uniform distribution between 1.1 and 1.3 for the alternative that was chosen in the RP, and between 0.7 and 0.9 for the others. The utility of the model used to build the SP choices was then given by:

$$U_{in}^{SP} = -1Time_{j\_n}^{SP} - 0.5Cost_{j\_n}^{SP} + \varepsilon_{jn}^{RP} + e_{jn}^{SP} , \quad [11]$$

where e_{jn}^{SP} is an exogenous quixotic error, build as random uniform between 0 and 1, that is only present in the SP choices.

Under this setting, the following four models were estimated:

- **True RP**: Estimated considering all the explanatory variables described in Eq. [9]. This model is used as a benchmark as the best possible outcome that could be obtained by the models that correct for endogeneity.

- **Uncorrected SP-off-RP**: SP model estimated neglecting the fact that the attributes are endogenous. This model is used as the benchmark of the do-nothing alternative.

- **SP-RP/CF**: Control-function method applied in the terms described in Section 3, using the RP attributes as instruments.

- **SP-RP/TW**: Full application of the T&W (3) method applied as explained in Section 2

These four models were analysed in terms of their finite sample properties. Because the correction of endogeneity in discrete-choice models produces a change in the scale of the estimators (9), we analyse the ratio of the estimators $\hat{\beta}_t / \hat{\beta}_c$ instead of the estimators themselves. The analysis was developed by repeating each experiment 100 times and analysing the sampling distribution of the estimators with a box-plot, as shown in Figure 1.

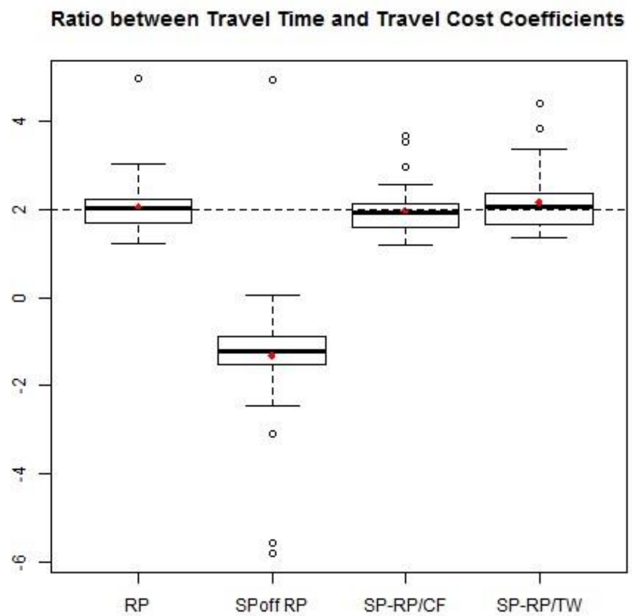

Figure 1: Boxplot for Different Methods to Correct For Endogeneity in the SP off RP Problem

The vertical axis in Figure 1 corresponds to the ratio $\hat{\beta}_t / \hat{\beta}_c$ and the methods are ordered in the horizontal axis. The dashed horizontal line in the middle marks the true population value of the ratio $\beta_t / \beta_c = -2$ and the boxplots are built from the estimators obtained from the 100 repetitions. The bold horizontal line on each boxplot corresponds to the median of the sampling distribution and the red triangle is the respective sample mean. The upper and lower box shows the 25% and 75% percentile, and the whiskers are limited by 1.5 the inter-quartile range.

Results show that, as expected, the RP model retrieves the population values quite well, while a model estimated on the SP-off-RP data with no correction shows a large negative bias in the estimation of the value of time. Both the proposed CF and the state of the art T&W method succeed in the correction of the finite sample bias since both are centred on the population value. Besides, it can be said also that T&W correction seems slightly more skewed up than the CF, since the sample mean and the median are less aligned and the whiskers are less symmetrical for the latter.

Regarding efficiency differences between the CF and T&W, despite T&W is more efficient in theory this does not show clearly in Figure 1. Actually, the spread of T&W´s sampling distribution is even slightly larger than the one of the CF method. This could be attributed to simulation error in T&W and to the fact that this synthetic data was created to have enough size and variance to properly estimate the problem. As we will show in Section 5, the contrary could occur if the data is poorer.

Finally, in terms of estimation times, we note that the RP model took 10.3 seconds in average to be estimated. The uncorrected SP off RP model took about the same, 12.3 seconds in average. In turn, the proposed CF correction takes about twice, 21.4 seconds, which is explained by the need for estimating an ordinary least square model in the first stage. Lastly, estimation with T&W method took 822 seconds in average, which is about 80 times the time needed for estimating the uncorrected model, highlighting the CF method is indeed not only easier to implement than T&W, but it is also significantly less onerous from a computational perspective. Some of the increase in estimation time for the T&W method comes from the joint estimation on both the RP and SP-off-RP data, but this is trivial compared to the requirement for numerical simulation.

5. APPLICATION WITH REAL DATA

For the analysis with real data, we revisit the freight choice experiment presented by Train and Wilson[1] (3). This looks at agricultural shippers in the Pacific Norwest, from the Whitman Country to Portland, using SP-off-RP. We revisit their dataset to illustrate in turn the application of the proposed CF method to the same problem.

The dataset consists of 103 shippers making the RP choice of route between the following 6 options, each one described by its rate (in dollars per ton), time (in days) and reliability (as % of arrival on time):

1. truck to Pasco and barge to Portland;
2. truck to another barge port and barge to Portland;
3. rail to Portland;
4. truck to a rail terminal and rail to Portland;
5. barge to Portland;
6. other.

Afterwards, the respondents faced three separate SP experiments, each one asking if they would shift their RP choice if the rate, time and reliability of their chosen alternative were, respectively, X percent higher, where X was randomly selected from a pre-defined set.

Table 1 summarizes the estimation results of the RP experiment, followed by each SP experiment separately and the pooled sample of the RP and the SP experiments accounting for the traditional RP/SP pooled sample, accounting for the scale change, but not for endogeneity (12). In all models, we use alternative 2 as the base alternative, with its ASC normalised to zero.

We take the RP results as the true values which do not suffer from endogeneity. We see a negative estimate for rate and time, where the latter is not statistically significant, along with a positive estimate for reliability. As a result of the small sample size, we see numerous problems with significance in the separate models for the three SP datasets. The jointly estimated RP/SP model performs better, although the coefficient of time is still not significantly different from zero, and indicates a lower scale for the SP data $\mu_{SP} < 1$.

[1] We would like to thank Kenneth Train for lending us the data used in his article.

Table 1. No Endogeneity Correction Portland SP-off-RP Freight Data

	RP		SP1 (Change cost)		SP2 (Change time)		SP3 (Change Relia)		Pooled RP/SP	
	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value
β_{Rate}	-0.125	-2.09	-0.07	-1.93	-0.12	-1.27	-0.13	-1.38	-0.175	-3.15
β_{Time}	-0.034	-0.948	-0.02	-0.48	0.00	0.20	-0.02	-0.49	-0.0417	-1.65
β_{Relia}	0.0322	3.26	0.01	1.30	0.01	1.46	0.01	0.77	0.03	3.52
ASC_1	-1.74	-3.14	-0.37	-0.92	-0.27	-0.65	-0.35	-0.85	-1.11	-2.70
ASC_2	-	-	-	-	-	-	-	-	-	-
ASC_3	1.075	1.853	-1.10	-1.77	-0.88	-1.51	-0.88	-1.65	0.102	0.241
ASC_4	-0.675	-1.344	-0.64	-1.60	-0.81	-1.94	-0.84	-1.94	-1.02	-2.64
ASC_5	-0.456	-0.674	-0.64	-0.85	-0.09	-0.10	-0.05	-0.06	-0.608	-1.03
ASC_6	-0.596	-0.563	-0.54	-0.35	-0.24	-0.16	-0.09	-0.05	-0.915	-0.96
μ_{RP}									1.00	-
μ_{SP}									0.428	3.80
Value of Time	0.274	0.878	0.216	0.44	-0.0389	-0.20	0.115	0.442	0.238	1.50
Value of Reliability	-0.257	-1.64	-0.189	-1.56	-0.115	-1.33	-0.0449	-0.952	-0.160	-2.16
N	103		81		82		84		350	
Mean LL per obs	-0.610		-0.843		-0.83		-0.838		-0.818	

Table 2 again reports the estimation results for the RP experiment, followed by each SP experiment separately and the pooled sample of the RP and the SP experiments, now accounting for endogeneity using the CF correction.

Table 2. CF Endogeneity Correction Portland SP-off-RP Freight Data

	RP		SP1 (Change cost) CF		SP2 (Change time) CF		SP3 (Change Relia) CF		Pooled RP/SP CF	
	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value
β_{Rate}	-0.125	-2.09	-0.10	-2.20	-0.11	-1.02	-0.13	-1.33	-0.175	-3.17
β_{Time}	-0.034	-0.948	-0.02	-0.57	-0.01	-0.33	-0.02	-0.51	-0.0439	-1.84
β_{Relia}	0.0322	3.26	0.02	1.36	0.01	1.34	0.01	0.86	0.0266	3.48
ASC_1	-1.74	-3.14	-0.43	-1.01	-0.35	-0.79	-0.38	-0.86	-1.04	-2.70
ASC_2	-	-	-	-	-	-	-	-	-	-
ASC_3	1.075	1.853	-1.10	-1.78	-0.83	-1.42	-0.85	-1.57	0.01	0.03
ASC_4	-0.675	-1.344	-0.76	-1.74	-0.89	-1.98	-0.85	-1.88	-1.04	-2.81
ASC_5	-0.456	-0.674	-1.27	-1.48	-0.38	-0.39	-0.10	-0.11	-0.84	-1.45
ASC_6	-0.596	-0.563	-0.70	-0.46	-0.58	-0.31	-0.16	-0.10	-0.96	-1.02
μ_{RP}									1.00	-
μ_{SP}									0.50	3.72
$\hat{\theta}_{cost}$			0.28	1.79					0.45	1.95
$\hat{\theta}_{time}$					0.46	1.84			0.94	1.96
$\hat{\theta}_{relia}$							-0.01	-0.34	-0.03	-0.86
Value of Time	0.274	0.878	0.19	0.544	0.08	0.297	0.12	0.456	0.25	1.65
Value of Reliability	-0.257	-1.64	-0.15	-1.32	-0.12	-1.10	-0.05	-1.10	-0.15	-2.08
N	103		81		82		84		350	
Mean LL per obs	-0.610		-0.82		-0.81		-0.84		-0.80	

Consider first the coefficients of the residuals $\hat{\theta}_{\cos t}$, $\hat{\theta}_{time}$ and $\hat{\theta}_{relia}$ for each SP separately and for the pooled sample. As noted by (7), the significance of this coefficient in the CF correction can be used as a test for the existence of endogeneity, which in this case means a test for the transfer of the RP error to the SP experiment. For the first two SP experiments the coefficient of the residuals $\hat{\theta}_{\cos t}$ and $\hat{\theta}_{time}$ are close to be significantly different from zero with 95% confidence. Although the critical value is not reached, lower values could be accepted in specification tests like this. In any case, this suggests that the endogeneity problem may not be too severe in this data and/or that it is masked in the poor variability and small sample size available in this case. $\hat{\theta}_{relia}$ for SP3 is clearly not significant, what can be attributed, with almost no doubt in this case, to the poor quality of the data, and we then also observe no improvement in fit compared to the model without the correction (unlike for the other games). For the pooled case, $\hat{\theta}_{\cos t}$ and $\hat{\theta}_{time}$ are even closer to the critical value to accept the presence of endogeneity, while $\hat{\theta}_{relia}$ remains not significant.

We finally report the results using the T&W correction approach on the same data. Our results differ from those in (3) as we allow for a non-unit RP error transfer by estimating ρ. The importance of doing this is confirmed by an improvement in model fit when comparing the model with ρ estimated to the original model. More importantly, we see that the VTT values obtained from the pooled models are substantially higher than the RP ones. However, while the estimate using the original approach shows a significant difference in the VTT between the RP value and the pooled result, this is no longer the case when allowing for $\rho \neq 1$. Some insights can also be gained when looking at the results for the separate SP games. These obtain more stable results than when working with the CF approach, but it is not clear whether the gains here are a result of a more efficient approach or the use of more data as the T&W approach is jointly using the SP and RP data, while the CF results are for SP alone.

One further observation to make relates to the SP scale estimates. We observe that for the individual SP games, the results for scale and error transfers are in line with the findings for the control function approach. This is no longer the case in the models pooling the data from different SP games, which surprisingly show a higher scale for SP than RP, contrary to earlier results. It seems that this is likely a direct result of using multiple SP games with the T&W approach creating correlations in the utilities between them through the error transfer from the RP data. We hypothesise that the approach thus potentially picks up effects that are not related to endogeneity bias but correlation across SP answers. In fact, if we take at face value the idea that the RP results are unbiased, then the results from a model using a correction should report VTT measures that are much closer to the RP ones than the T&W approach. To test this hypothesis, we also estimated a version of the model which uses independent but identically distributed error transfers from the RP to the three different SP games. These results are shown in the final column of Table 3. They produce a much lower log-likelihood, showing that the earlier gains were at least in part a result of capturing correlation across games. Crucially, they also show VTT results that are in line with the CF method and much closer to the RP results.

Table 3. T&W Endogeneity Correction Portland SP-off-RP Freight Data

	SP1 (Change cost) T&W		SP2 (Change time) T&W		SP3 (Change Relia) T&W		Pooled RP/SP T&W with ρ		Pooled RP/SP (Train & Wilson, 3)		Pooled RP/SP T&W with ρ but independent draws	
	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value
β_{Rate}	-0.2117	-0.89	-0.1561	-1.38	-0.1705	-1.4	-0.2435	-3.04	-0.2047	-2.45	-0.2087	-1.26
β_{Time}	-0.0373	-1.11	-0.0383	-1.15	-0.0388	-1.19	-0.1355	-3.27	-0.1454	-4.13	-0.0444	-1.28
β_{Relia}	0.0336	2.78	0.0334	2.8	0.0345	2.78	0.0288	4.19	0.0278	3.67	0.0316	2.17
ASC_1	-1.4505	-1.77	-1.5627	-2.24	-1.5229	-2.3	-0.4337	-0.85	-0.126	-0.3	-1.1003	-1.55
ASC_2	-	-	-	-	-	-	-	-	0	-	0.0000	NA
ASC_3	0.8259	1.67	0.8885	1.75	0.8611	1.72	1.625	3.07	1.0269	2.2	0.2273	0.29
ASC_4	-0.7284	-2.11	-0.7827	-2.23	-0.7868	-2.44	-0.0443	-0.12	-0.1227	-0.37	-0.9687	-2.67
ASC_5	-0.8838	-0.61	-0.5186	-0.43	-0.5192	-0.41	-0.8033	-0.63	-0.9303	-0.79	-0.7015	-0.48
ASC_6	-1.2553	-0.43	-0.9069	-0.5	-1.0366	-0.52	-1.2207	-0.67	-0.9287	-0.59	-1.2378	-0.46
μ_{RP}							1.00	-	1.00	-	1.00	1
μ_{SP}	0.3068	1.01	0.2733	1.85	0.406	2.09	3.3763	1.7	5.6575	2.31	0.4050	1.46
ρ	0.8051	0.71	0.4437	0.62	0.7053	1.31	1.3399	4.75	1	-	0.3452	0.5
Value of Time	0.176	0.69	0.245	0.90	0.228	0.87	0.557	3.00	0.7103	2.84	0.213	0.86
Value of Reliability	-0.159	0.89	-0.214	1.20	-0.202	1.38	-0.118	8.62	-0.136	4.33	-0.152	1.73
N	184		185		187		350		350		350	
Mean LL per obs	-0.749		-0.744		-0.741		-0.666		-0.690		-0.816	

6. CONCLUSION

Stated Preference (SP) data allows modelling choices under a controlled experimental environment and at low cost, but suffers from cognitive dissonance between the stated and the actual choice behavior. To lessen such a drawback, it is common practice to build SP attributes and alternatives, from observed Revealed Preference (RP) choices. For example, in SP-off-RP data, the RP chosen alternative is worsened and other alternatives are improved to induce a choice change. This practice substantially increases the realism of the choice task, compared to SP settings that ignore the RP data, but with the undesired cost of introducing endogeneity. This source of endogeneity was largely neglected in literature and practice until (3) proposed a FIML solution that can be implemented with simulation. This seminal work has been profusely cited but rarely applied, arguably because of its various practical difficulties.

In this article we propose a LIML approach to address the SP-off-RP problem using a method that does not need simulation, can be applied with canned software and uses data that is already available for the stated problem. The proposed method can be seen as an application of the Control-Function (CF) method to correct for endogeneity in discrete choice models, using the RP attributes as instrumental variables.

We discuss the theoretical and practical advantages and disadvantages of the CF and T&W methods and illustrate them using Monte Carlo and real data. Results show that both methods can successfully address the endogeneity problem that arise from the use of SP-off-RP data, and that the proposed CF method is indeed significantly easier to implement and less onerous from a computational perspective. The T&W method proved to be more efficient, which could matter when the data available is too poor, however, there are also concerns that the results are driven by accounting for correlation across SP choices.

ACKNOWLEDGEMENT

Stephane Hess was supported by the European Research Council through the consolidator grant 615596-DECISIONS.

REFERENCES

1. Hensher, D. A. (2010). Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B: Methodological, 44(6), 735-752.
2. Shinghal, N., & Fowkes, T. (2002). Freight mode choice and adaptive stated preferences. Transportation Research Part E: Logistics and Transportation Review, 38(5), 367-378.
3. Train, K., & Wilson, W. W. (2008). Estimation on stated-preference experiments constructed from revealed-preference choices. Transportation Research Part B: Methodological, 42(3), 191-203.
4. Train, K. E., & Wilson, W. W. (2009). Monte Carlo analysis of SP-off-RP data. *Journal of Choice Modelling*, *2*(1), 101-117.

5. Van Cranenburgh, S., Chorus, C. G., & Van Wee, B. (2014). Vacation behaviour under high travel cost conditions–A stated preference of revealed preference approach. *Tourism Management*, *43*, 105-118.
6. Heckman, J. (1978), “Dummy Endogenous Variables in a Simultaneous Equation System,” Econometrica, 46, 931-959.
7. Rivers, D. and Q. Vuong (1988), “Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models,” Journal of Econometrics, 39, 347-366.
8. Petrin, A. and K. Train (2002), “Omitted Product Attributes in Discrete Choice Models,” Working Paper, Department of Economics, University of California, Berkeley, CA.
9. Guevara, C. A., & Ben-Akiva, M. E. (2012). Change of scale and forecasting with the control-function method in logit models. Transportation Science, 46(3), 425-437.
10. Guevara, C., & Ben-Akiva, M. (2006). Endogeneity in residential location choice models. Transportation Research Record: Journal of the Transportation Research Board, (1977), 60-66.
11. Karaca-Mandic, P., & Train, K. (2003). Standard error correction in two-stage estimation with nested samples. The Econometrics Journal, 6(2), 401-407.
12. Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., & Rao, V. (1994). Combining revealed and stated preferences data. Marketing Letters, 5(4), 335-349.