

## ***Research note: Strategic behaviour in multiple purpose data collection – a travel cost application to Barcelona Zoo***

VERÓNICA FARRERAS

*Argentine Institute of Nivology, Glaciology and Environmental Sciences (IANIGLA – CONICET), 5500 Mendoza, Argentina. E-mail: vfarreras@mendoza-conicet.gob.ar.  
(Corresponding author.)*

PERE RIERA

*ICTA, Autonomas University of Barcelona, E-08193 Bellaterra, Spain.  
E-mail: prieram@gmail.com.*

Some tourism and recreational values are estimated by the travel cost method (TCM) using secondary data. The purpose of the primary collection could suffer from incentive compatibility problems. This being the case, the TCM value estimates may be biased, and extra care should be taken. This article provides theoretical background and illustrates the problem with a case study of visitors to Spain's Barcelona Zoo.

*Keywords:* incentive compatibility problems; secondary data; tourism values; recreational values; travel cost method

Recreation is one of many services present in today's leisure activities. The value that users attach to recreation may be substantial, although, to some extent, it is not reflected by market prices and is provided as a public good. Valuing recreational use requires the monetary valuation of a non-market good. Although a number of techniques are available to value non-market goods, the one most widely used to measure recreational values is the travel cost method (TCM). The TCM relies on data of passed events reported by individuals. Sometimes these data are collected from secondary data sources (Willis and Benson, 1988; Bell and Leeworthy, 1990).

Since individuals often rely on their memory to provide data about solicited events, discrepancies detected between actual and reported past events have been mainly attributed to a recall problem or recall bias (Gems *et al.*, 1982; Westat, 1989). However, there has been, to our knowledge, no attempt to explain this from an economics perspective. This paper claims that, in given instances, the purpose of the primary collection could suffer from incentive compatibility problems.<sup>1</sup> Therefore, individuals may in fact have the possibility of being better off by responding something other than the truth. For instance, if an

individual is asked how many times she or he went fishing for a particular species in the last year, and reports a higher number, the administration might devote more money to reintroduce the species. By doing so, the respondent might benefit more than if she or he had reported the actual number of fishing trips motivated by this particular species. If then the administration uses the number of visits reported by fishermen to estimate the recreational value of the fishing area through the TCM, the incentives problem may be transferred to the recreational values. This being the case, the TCM value estimates may be biased, and extra care should be taken. This paper provides theoretical background and illustrates the problem through a case study of visitors to Barcelona Zoo, Spain.

The rest of the paper is organized as follows. The next section describes the theoretical background. The subsequent section deals with an empirical example involving the reporting of the number of visits to Barcelona Zoo. In the penultimate section we discuss the results and in the final section we present our conclusions.

### Theoretical background

A social choice mechanism,  $\tau = (M_1, \dots, M_n, \dots, M_N, g(\cdot))$ , can be defined as a collection of  $N$  messages (or sets of messages)  $(M_1, \dots, M_n, \dots, M_N)$  given by  $N$  agents  $(1, 2, \dots, n, \dots, N)$ , with  $N \geq 2$ , and a function  $g: M_1 \times \dots \times M_n \times \dots \times M_N \rightarrow X$ , where  $M_1 \times \dots \times M_n \times \dots \times M_N$  denote the message space, and  $X$  the set of at least two social decisions under consideration or social outcomes. Following the usual notation,  $M_{-n}$  denotes the set of messages given by all the agents but agent  $n$  (Villa and Manrique, 2003).

From the perspective of agent  $n$ , she or he is asked to provide a message knowing that a rule  $g(\cdot)$  will be applied according to the messages from all the agents to select an element of  $X$ . This means that both  $g(\cdot)$  and  $X$  are publicly known. To provide the message, each agent observes a parameter  $\theta_n$ , often called *type*, which represents the preferences of the agent over the possible outcomes. The preferences of the  $N$  agents can be represented by a set  $(\theta_1, \dots, \theta_n, \dots, \theta_N)$ , where  $\theta_1 \times \dots \times \theta_n \times \dots \times \theta_N$  is the space of the preferences of all the agents. Function  $\Theta: \theta_1 \times \dots \times \theta_n \times \dots \times \theta_N \rightarrow (M_1, \dots, M_n, \dots, M_N)$  relates preferences with messages, including the possible strategies to follow. It is assumed that types are privately known. This implies that each agent ignores the exact preferences of the rest of the agents, and that also the social planner ignores the preferences of the agents. Therefore, from the planner's perspective, she or he obtains information on agents' preferences by collecting messages from them, and according to those messages applies a decision rule  $g(\cdot)$  to select an element of  $X$  for its implementation (Green and Laffont, 1977; Campbell, 1995; Corchón, 1996).

A mechanism is defined as manipulable when there is at least one agent,  $n$ , who, given the messages of the rest of the agents, with a false message  $(M_{-n}, M'_n)$  could obtain a social decision  $x \in X$  that is more favourable, according to her or his preferences  $\theta_n$ , than the one that would have obtained by providing the true message  $(M_{-n}, M_n)$  (Villa and Manrique, 2003). A manipulable mechanism is also called not incentive compatible.

### *A manipulable mechanism for policy decisions*

Depending on its design, a mechanism applied to decide on a given policy, for example whether to implement subsidies, may be manipulable (Green and Laffont, 1977). A description of such a mechanism follows. The mechanism corresponds to the application example that will be dealt with later in this paper. Suppose a planner demanding a message from  $N$  rational economic agents by means of a survey. The message consists in a natural number reflecting the number of trips made in a given period of time to a certain zoo. After applying the decision rule to the messages collected, the planner decides whether to subsidize the consumption of the good (for example, introducing a discount scheme to the zoo entrance fee) and, if implemented, defines the amount of the discount. The amount of discount one particular agent receives is the same as the one received by every other eligible agent, regardless of the message given by the agent.

Thus, the social outcomes or elements (at least two) of set  $X$ , are  $(x_1, x_2, \dots, x_j, \dots, x_J)$  where the elements in parentheses denote the possible values of the subsidy ordered according to its amount. The zoo example,  $x_1$  denotes no discount (that is, a discount of 0 monetary units),  $x_2$  corresponds to a strictly positive discount, larger than  $x_1$ , and so on, being  $x_j$  the largest discount in the list of social outcomes contemplated (in the example,  $J = 3$ ). The elements of  $X$  correspond to real positive numbers, such that  $x_1 < x_2 < \dots < x_j < \dots < x_J$ . The welfare of each agent increases as the subvention rises,  $x_j$ .

Consider now a set  $P = (P_1, P_2, \dots, P_j, \dots, P_J)$ , such that each element of  $P$  is a real positive number which value is fixed and known when conducting the survey, with  $P_1 < P_2 < \dots < P_j < \dots < P_J$ . Additionally, suppose that the decision rule  $g(\bullet)$  implies first the estimation of an average,  $V$ , of trips or visits that  $N$  agents state to have undertaken in a given period of time. Next, if  $V \leq P_1$ , the outcome to be selected by the planner would be  $x_1$ . In general, if  $P_{j-1} < V \leq P_j$ , the selected outcome would be  $x_j$ . Notice that the computation of  $V$  depends on the messages provided by each agent, that is, the number of visits to the zoo in a given period of time, in the case example. Reporting a higher number of trips results in a higher average, while understating them lowers the value of  $V$ .

To demonstrate that this example of a mechanism is not incentive compatible, it suffices to show one case where manipulation may occur. For this purpose, consider the case of a combination of values of  $V$ ,  $M_{-n}$ ,  $M_n$  and  $P_1$ , with agent  $n$  preferring a higher subsidy to a lower one, such that  $V$  exactly equals  $P_1$ . Therefore, the outcome to be selected by the planner would be  $x_1$  in contrast to the preferences of the agent  $n$ . Then, if agent  $n$  states a higher number of trips,  $M'_n$ , than those really undertaken,  $M_n$ , the value of  $V$  will exceed  $P_1$ , since  $V(M_{-n}, M'_n) > V(M_{-n}, M_n)$ , and the outcome would change from no discount (in the example) to a strictly positive discount, benefiting agent  $n$ . The mechanism is, therefore, not incentive compatible.

### **Empirical application**

During the last quarter of 2003 and first of 2004, face-to-face *in situ* interviews with a random sample of visitors exiting Barcelona Zoo were conducted in the

context of a travel cost exercise. Most of the questionnaire followed a standard pattern, but one question required to conduct the travel cost exercise was prepared in relation to the incentive compatibility problem. More specifically, before requesting the number of trips made to zoo in the past, we stated a purpose other than measuring the recreational value of the zoo. Accordingly, half of the sample (group B, or incentive group) faced a version of the questionnaire that informed visitors residing in the Barcelona province that the provincial authority was considering the introduction of an entrance fee discount to their future visits. The amount of the discount, which if implemented would be the same for every resident, was to be determined according to the following rule. No discount if, according to the survey results, the mean of the residents' visits over the last 5 years was two or fewer visits; €3 discount if the average were more than two and fewer than five visits; and €7 if the average were five or more visits. This can be seen as a manipulable mechanism, according to the example discussed in the previous section. The other half of the sample (group A or control group) as well as visitors not residing in Barcelona province, were not given this information, so no purpose was stated before requesting the number of visits to the zoo in the last 5 years.

The total number of interviews was 167,114 with Barcelona province residents, while the remaining 53 were with non-residents. The sample composition of gender and age agreed with the visitor figures published by the zoo administration (Direcció de Comunicació Corporativa i Qualitat, 2001).

## Hypotheses and results

The empirical application is organized into two stages. First, we present the incentive compatibility hypothesis and the results associated to the entrance fee discount scheme. We then test the hypothesis related to the transfer of the incentives problem to the recreational values estimated by the TCM.

### *Entrance fee discount exercise*

The null hypothesis ( $H_0$ ) states that the messages reported by the respondents are independent of the economic incentives provided by the information on the entrance fee discounts.<sup>2</sup> More specifically, that the number of visitors from group A reporting two or fewer visits, three or four, and five or more visits in the last 5 years would be similar to the numbers observed from group B once compensated by differences in the size of the subsamples. The alternative hypothesis ( $H_1$ ) implies dependency on the information given in the questionnaire. That is, some respondents react according to the discount incentives, and so, the number of visitors stating any given number of visits would be different between groups A and B. The non-rejection of  $H_0$  would suggest that the expected strategic behaviour is not confirmed, whereas its rejection would not discard such behaviour.<sup>3</sup>

Table 1 shows the number of messages for the different combinations of visits intervals and groups of economic agents. As shown, fewer individuals from B reported two or fewer visits, and more reported three or four and five or more. The chi-squared statistic value is 7.81, while the critical value for a 95%

Table 1. Data and test results for  $H_0$ .

Visit intervals	Control group (A)	Incentive group (B)
2 or fewer visits	22	7
3–4 visits	18	21
5 or more visits	21	25
Observations	61	53
$\chi^2$		7.81
$p$ -value		0.02

confidence level with 2 degrees of freedom is 5.99, which suggests the rejection of the first null hypothesis at 5% significance level. In other words, the notion cannot be discarded that individuals belonging to the incentive group B were more likely to engage in strategic behaviour. This result is as expected, since the mechanism is manipulable as underlined in the theoretical section.<sup>4</sup>

#### *TCM exercise*

It is evident that the purpose of the primary collection suffers from incentive compatibility problems (Table 1). Some individuals perceived that they might be better off by misreporting the datum they were asked to state, and acted accordingly. Therefore, since the mechanism of the primary collection is not incentive compatible, we hypothesize that the TCM value estimates may also be biased owing to strategic behaviour.

Correspondingly, the null hypothesis,  $H_0$ , could be formulated as follows. The data misreported in primary collection does not result in higher TCM value estimates for individuals belonging to group B. The alternative hypothesis,  $H_1$ , implies that the TCM value estimates for individuals belonging to group B is higher than that estimated for group A. Therefore, TCM value estimates are biased due to strategic behaviour.

Formally,

$$H_0: E_B - E_A = 0$$

$$H_1: E_B - E_A > 0, \quad (1)$$

where  $E$  is the consumer surplus per visit estimated by the TCM; and  $A$  and  $B$  stand for, respectively, visitors residing in Barcelona province who did not have the information about possible future entrance fee discounts (group A) and non-resident visitors. The other subsample consisted of visitors informed about the entrance fee discount policy (group B) and visitors not residing in Barcelona province.

The non-rejection of  $H_0$  would suggest that the TCM value estimates using secondary data do not result in biased values although the mechanism of the primary collection is manipulable. To test these hypotheses we use the non-parametric statistic proposed by Poe *et al* (1997).

Table 2 shows the results of the travel cost exercise for both groups. We

Table 2. TCM results.

Variable	Control group (A)	Incentive group (B)
<i>Parameters of Poisson probability</i>		
Constant	2.256* (10.748)	1.889* (7.798)
Travel cost	0.052* (-6.292)	-0.032* (-3.286)
<i>Standard deviation of heterogeneity</i>		
$\sigma(\epsilon)$	0.992* (10.012)	0.743* (8.529)
Log-L	-275.885	-251.726
Observation	114	106

Note: *t*-statistics in parentheses; \*statistically significant at the 99% confidence level.

Table 3. Estimated consumer's surplus per visit.

Control group (A)	Incentive group (B)	Significance level ( $E_B - E_A$ )
19.215 (14.579, 27.861)*	31.639 (19.933, 84.162)*	0.044

Note: \*95% confidence interval.

model the visits to Barcelona Zoo following a lognormal model (Winkelmann, 2004). The dependent variable gathers the number of visits to the zoo during the last 5 years. The exogenous variable is the travel cost, including the entrance fee and the round trip travel expenses. The negative sign of the estimated travel cost coefficient is as expected and statistically significant at 99% confidence level. It shows that the probability of an additional visit to the zoo decreases as the travel costs rise. The significant *t*-statistic for  $\sigma(\epsilon)$  indicates the existence of heterogeneity. This result supports the use of the lognormal model which adequately accounts for the latent heterogeneity.

Table 3 shows each group's estimated consumer surplus per visit,  $E_A$  and  $E_B$ . The consumer surplus per visit can be inferred by calculating the ratio  $-1/B_{TC}$ , where  $B_{TC}$  is the regression coefficient of the travel cost (Creel and Loomis, 1990). As shown in the first row, group B's consumer surplus per visit is €12.43 higher than group A's consumer surplus. The second row shows the confidence intervals for the estimated consumer surplus, which were calculated based on 1,000 random draws (see Krinsky and Robb, 1986).

The last column of Table 3 tests the  $H_0$  from Equation (1) – the equivalence between  $E_A$  and  $E_B$ . The consumer surpluses estimated by TCM,  $E_A$  and  $E_B$ , are not equivalent at the 5% significance level. More specifically, the consumer surplus per visit of group B is statistically higher than that of group A. This result is in line with the previous exercise of the entrance fee discounts. Therefore, it cannot be discarded that the recreational values estimated by TCM

using secondary data may be biased when the purpose of the primary collection suffers from incentive compatibility problems.

This problem may happen in any of the different models of the TCM – individual, hedonic and zonal – and specifications determining the behaviourally relevant cost of travel – researcher-assigned costs and subjective costs. The reason is that the TCM relies on data reported by individuals. And, if the purpose of the primary collection suffers from incentive compatibility problems, the TCM value estimates may be biased. Hence, the incentives problem arises from data collection regardless the model of the TCM and specification of travel costs.

## Conclusion

Traditionally, strategic behaviour in the valuation of public goods has not been associated to the TCM. However, when the messages are elicited in the context of a social choice mechanism, the mechanism may not be incentive compatible, and therefore individuals may misreport on messages because of incentives to behave strategically.

This paper shows how an incentives problem could be transferred to the recreational values estimated by the TCM when the purpose of the primary collection is not incentive compatible. Barcelona Zoo illustrates a case where an incentives problem could be a possible complementary explanation to the traditional recall problem. Therefore, decision makers relying on TCM value estimates as a component in their decision making must take extra steps to evaluate the reliability of these values since the purpose of data collection may be manipulable.

These findings reinforce the importance to identify, reduce, and manage the incentives problem that could arise from data collection. An adequate assessment of how the information was gathered, analysed and presented ought to be seen as an integral part of the good practice in techniques to value non-market goods.

## Endnotes

1. As the one commonly studied in the literature of social choice or implementation (Green and Laffont, 1977; Campbell, 1995; Corchón, 1996).
2. The usual Pearson's chi-square statistic is used to test the null hypothesis of independency.
3. When a mechanism is manipulable, and economic theory may predict that agents engaged in it will behave accordingly, empirical evidence suggests that reactions to incentives and other types of verbal cues in economic surveys are heterogeneous across individuals and situations (Nunes, 2002; McElroy and Seta, 2003). Therefore, testing for statistically significant differences in the stated number of previous visits is a key test to empirically demonstrate the basics of the theoretical section.
4. Besides the strategic behaviour problem associated to the entrance fee discount scheme in questionnaire B, respondents from both A and B may still face other incentives and memory problems in recalling the number of times they went to the zoo in the last 5 years. However, those other incentives and memory problems must be present in A as well as in B.

## References

- Bell, F., and Leeworthy, V. (1990), 'Recreational demand by tourists for saltwater beach days', *Journal of Environmental Economics and Management*, Vol 18, No 3, pp 189–205.



- Campbell, D. (1995), *Incentives: Motivation and the Economics of Information*, Cambridge University Press, Cambridge.
- Corchón, L. (1996), *The Theory of Implementation of Socially Optimal Decisions in Economics*, Macmillan, London.
- Creel, M., and Loomis, J. (1990), 'Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in California', *American Journal and Agricultural Economics*, Vol 72, No 2, pp 434–441.
- Direcció de Comunicació Corporativa i Qualitat (2001), *Enquesta Visitants del Parc Zoològic de Barcelona*, Ajuntament de Barcelona, Barcelona.
- Gems, B., Ghosh, D., and Hitlin, R. (1982), 'A recall experiment: impact of time on recall of recreational fishing trips', *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp 372–375.
- Green, J., and Laffont, J. (1977), 'Characterization of satisfactory mechanism for the revelation of preferences for public goods', *Econometrica*, Vol 45, No 2, pp 727–738.
- Krinsky, I., and Robb, L. (1986), 'On approximating the statistical properties of elasticities', *The Review of Economics and Statistics*, Vol 68, No 4, pp 715–719.
- McElroy, T., and Seta, J. (2003), 'Framing effects: an analytic-holistic perspective', *Journal of Experimental Social Psychology*, Vol 39, No 6, pp 610–617.
- Nunes, P. (2002), *The Contingent Valuation of National Parks: Assessing the Warmglow Propensity Factor*, Edward Elgar, Cheltenham.
- Poe, G., Welsh, M., and Champ, P. (1997), 'Measuring the difference in mean willingness to pay when dichotomous choice contingent valuation responses are not independent', *Land Economics*, Vol 73, No 2, pp 255–267.
- Villa, E., and Manrique, O. (2003), *Teoría de la Implantación y Diseño de Mecanismos*, Universidad Externado de Colombia, Bogotá.
- Westat (1989), 'Investigation of possible recall/reference period bias in national surveys of fishing, hunting, and wildlife-associated recreation', Westat, Inc. for the US Department of the Interior, Fish and Wildlife Service, Washington, DC.
- Willis, K., and Benson, J. (1988), 'A comparison of user benefits and costs of nature conservation at three nature reserves', *Regional Studies*, Vol 22, No 5, pp 417–428.
- Winkelmann, R. (2004), 'Health care reform and the number of doctor visits – an econometric analysis', *Journal of Applied Econometrics*, Vol 19, No 4, pp 455–472.