Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Transportation Research Part A 40 (2006) 1–18

TRANSPORTATION
RESEARCH
PART A

www.elsevier.com/locate/tra

On fitting mode specific constants in the presence of new options in RP/SP models

Elisabetta Cherchi [a,1], Juan de Dios Ortúzar [b,*]

[a] CRiMM, Dipartimento di Ingegneria del Territorio, Facoltà di Ingegneria, Università di Cagliari, Piazza d'Armi, 16, 09123 Cagliari, Italy
[b] Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile, Casilla 306, Cod. 105, Santiago 22, Chile

Received 24 January 2003; received in revised form 30 March 2005; accepted 4 April 2005

Abstract

We treat the problem of fitting alternative specific constants (ASC) in models estimated with a mixture of revealed preference (RP) and stated preference (SP) data to forecast the market shares of new alternatives. This important problem can have non-trivial solutions, particularly when some of the SP alternatives are completely revamped versions of existing ones (i.e., an advanced passenger train replacing a normal railway service). As there is no explicit treatment of this problem in the literature we examined it in depth and illustrated it empirically using data especially collected to analyse mode choice in a corridor to the West of Cagliari. We propose a hopefully useful guide to this art (as no practical recipes seem to serve all purposes). Careful specification of the systematic component of utility functions in RP and SP, including the ASC, serves to illuminate the true nature of the underlying error structure in the different data sets, yielding superior forecasting models.

Keywords: RP/SP models; Model specification; Error structure; Forecasting new options

* Corresponding author. Tel.: +56 2 686 4822; fax: +56 2 553 0281.
E-mail addresses: echerchi@unica.it (E. Cherchi), jos@ing.puc.cl (J. de Dios Ortúzar).
[1] Tel.: +39 70 675 5274; fax: +39 70 675 5261.

0965-8564/$ - see front matter
doi:10.1016/j.tra.2005.04.002

1. Introduction

The recommended approach to estimating choice models intended to forecast demand for new alternatives involves pooling (if available) revealed preference (RP) and stated preference (SP) data. The former are based on observations of actual choices and allow to characterise current travel behaviour, while the latter provide information about user preferences for new alternatives or for alternatives that differ radically from existing ones. The combined use of both types of data allow to exploit their respective advantages and to overcome their specific limitations (Ben-Akiva and Morikawa, 1990; Bradley and Daly, 1997; Louviere et al., 2000).

In joint estimation some attributes may be specified as generic for both the RP and SP alternatives and others specific to each subset. Variable specification depends on the analyst's choice; this is based on the type of data containing the variable, where it is measured with greater precision, and on the estimation results. Since alternative specific constants (ASC) allow simple logit (MNL) models to reproduce observed market shares, their specification as generic for the two data sources depends on the specific application context and on the market the analyst is trying to reproduce. Also, since the ASC represent the mean of the error difference and the errors are strictly related to the systematic utility specification, their values and significance are influenced by the degree to which the model is able to reproduce real behaviour. Thus, quite different results may be obtained depending on the way the RP/SP ASC are specified and it is not straightforward to define the best specification. As far as the authors are aware, this issue has not been addressed directly elsewhere.

The rest of the paper is organised as follows. Section 2 analyses some theoretical aspects concerning the inclusion of ASC in joint RP/SP model estimation, with particular reference to their role in the random and systematic utility components and to the MNL property of reproducing actual market shares. Section 3 briefly describes the data used to estimate the models presented in Section 4, where many different ASC specifications are tested and analysed, with reference to various hypotheses about the specification of the variables and error terms. Section 5 summarises our main conclusion.

2. Alternative-specific constants in logit models

Following the classical formulation of random utility models (Domencich and McFadden, 1975), individuals are assumed to choose among several available alternatives ($A_j \in \underline{A}(q)$, where $\underline{A}(q)$ is the set of options available to individual q), associating to each an index of preference (called utility) that depends on its characteristics and those of the individual:

$$U_{qj} = U(\underline{X}_{qj}, \varepsilon_{qj}), \tag{1}$$

where $\underline{X}_{qj}$ is a vector of measurable attributes, and ε_{qj} is a random term typically introduced by the modeller due to his/her inability to obtain perfect information about the individuals. Under the almost universally accepted assumption of additive random terms, utility is specified as the sum of a systematic, representative or observable part (V_{qj}) and a random error (ε_{qj}):

$$U_{qj} = V_{qj}(\underline{X}_{qj}) + \varepsilon_{qj}. \tag{2}$$