The Journal of Choice Modelling 5 (2012) 145–156

Contents lists available at SciVerse ScienceDirect

The Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

Latent variables in discrete choice experiments ☆

Cam M. Rungie [a,*], Leonard V. Coote [b,1], Jordan J. Louviere [c,2]

[a] School of Marketing, University of South Australia, Adelaide, SA 5000, Australia
[b] UQ Business School, University of Queensland, Brisbane, Qld. 4072, Australia
[c] Centre for the Study of Choice, University of Technology, Sydney, Ultimo, NSW 2007, Australia

ARTICLE INFO

Available online 20 May 2013

Keywords:
Structural equations
Latent variables
Discrete choice experiments

ABSTRACT

This paper describes and applies a general approach for incorporating factors with structural equations into models for discrete choice. The approach gives form to the covariance matrix in random coefficient models. The factors act directly on the random coefficients as unobserved attributes. The structural equations allow the factors to act on each other building structures that can represent a variety of concepts such as global heterogeneity and segmentation. The practical outcomes include parsimonious and identified models with rich covariances and better fit. Of greater interest is the ability to specify models that represent and test theory on the relationships between the taste heterogeneities for covariates and in particular between the attributes within a discrete choice experiment. The paper describes the general model and then applies it to a discrete choice experiment with seven attributes. Four competing specifications are evaluated, which demonstrates the ability of the model to be identified and parsimonious. The four specifications also demonstrate how competing *a priori* knowledge of the structure of the attributes used in the experiment can be empirically tested and evaluated. The outcomes include new behavioral insights and knowledge about choice and choice processes for the subject area of discrete choice experiments.

1. Introduction

Random utility theory assumes that the utility individuals derive from a choice object can be partitioned into a systematic component, capturing the attributes of the choice alternatives and the characteristics of the individual decision makers, and a random idiosyncratic component. Building on this foundation, McFadden (1974, 2001) derived the conditional logit model to represent discrete choice, extending the work of Thurstone (1927), Luce (1959), and Marschak (1960). Many different developments in choice models have been evident since, with the motivation for much of this development being to better represent choice processes (Ben-Akiva and Lerman, 1985; Hensher et al., 2005). For example, Train (2003) proposed a random coefficients model where utility parameters vary over individuals with distributions reflecting some latent choice process(es). The systematic component of utility is a weighted sum of the covariates, representing the attributes and characteristics, where the weights are random coefficients (McFadden and Train, 2000). However, generating

☆ This research was supported by a grant from the Australian Research Council (DP0880074). The authors gratefully acknowledge seminar participants at the Centre for the Study of Choice, University of Technology Sydney, Australia for their comments on earlier drafts of this paper.

* Corresponding author. Tel.: +61 8 8302 0768; fax: +61 8 8302 0765.
E-mail addresses: cam.rungie@unisa.edu.au (C.M. Rungie), l.coote@business.uq.edu.au (L.V. Coote), jordan.louviere@uts.edu.au (J.J. Louviere).
[1] Tel.: +61 7 3365 9721; fax: +61 7 3365 6988.
[2] Tel.: +61 2 9514 3993; fax: +61 2 9514 3535.

1755-5345
http://dx.doi.org/10.1016/j.jocm.2013.03.002

valid, parsimonious, and identified specifications for the covariance matrix, $\sum$, for the random coefficients has been problematic. In this paper we describe an approach using factors with structural equations that is shown to be effective. The approach, known as structural choice modeling (SCM), has practical benefits in generating identified and parsimonious models that fit the data better. It has the even more important property of being able to validly represent and test propositions and *a priori* knowledge regarding the form of $\sum$ and the structure of the taste heterogeneity for the covariates. The purpose of this paper is to describe the use and benefits of incorporating factors and structural equations into models for choice data.

The paper starts with a review of the literature on specifying the form of $\sum$ in random utility models. It develops and describes SCM. The SCM approach is used to analyze a discrete choice experiment that investigates postgraduate candidates' preferences for postdoctoral employment in universities. Four competing specifications of the model are fitted to the data and tested.

2. Literature

Factors have been incorporated explicitly into choice models to represent unobserved attributes of objects (often brands) of interest. For example, Elrod and Keane (1995) use a factor analytic probit model to study consumers' repeated brand purchases with model parameters representing market structure. That is, the factors are assumed to be unobserved attributes of brands and the associated factor loadings capture the location of each brand on these latent attributes. Taste heterogeneity for the brands is modeled as having several dimensions, one for each unobserved attribute. Each loading shows the contribution of the dimension to the taste heterogeneity of the brand. The loadings collectively show the positioning of the brand relative to the unobserved attributes. Keane (1997) extends the approach to a range of attributes and to the analysis of state dependency. Typically, the aforementioned models have been applied to observational data (i.e., revealed preference data that reflects actual choices of decision makers in real markets).

By contrast, stated preference methods, such as the discrete choice experiments (DCE) discussed by Louviere and Woodworth (1983) and Louviere et al. (2000), control for colinearity and incidental correlations prevalent in observational data, enabling evaluation of the causal impact of attributes on choice. We focus on modeling attributes in DCEs in part because of these aspects of DCE data. Models with factors also are used to combine data sets from the same individuals and to combine RP (revealed preference, or real market choices) and SP (stated preference) data (Ben-Akiva and Morikawa, 1990; Hensher et al., 1999; Louviere et al., 1999, 2002; Ben-Akiva et al., 2002; Morikawa et al., 2002). Extending the work by applying SCM to combining data sets from DCEs has considerable potential, such as combining SP and RP. DCEs with the same individuals but different combinations of attributes, and even with different choice tasks, can be modeled simultaneously and the common attributes linked (Rungie et al., 2011).

Building on the models of Ben-Akiva, Keane and others, Walker (2001) developed a general model combining factors and structural equations. The random coefficients for the covariates are functions of factors, with the factors also being functions of each other. This structural choice model is the basis for our work and is described in greater detail in the next section. However, we first consider some other important developments in the use of factors and structural equations in discrete choice.

Choice models proposed by Maydeu-Olivares and Böckenholt (2005) can achieve similar outcomes, but use a different data generation and modeling approach. Maydeu-Olivares and Böckenholt use data from paired comparison and ranking tasks that yield binary measures of stated preferences for the brands they studied. The binary choices are transformed using threshold models, with the transformed data analyzed with widely available standard structural equation modeling software (Jöreskog, 1970, 1973; Bollen, 1989; Jöreskog and Sörbom, 1996). This approach allows one to specify the form of the sample covariance matrix, particularly for the choice models Thurstone (1927) proposed. In principle, it can be extended to more complex choice designs, but such extensions cannot be easily accommodated with existing SEM software. Furthermore, with the exception of Bayesian approaches (e.g., Dunson, 2000; Lee, 2007), using threshold models to transform binary responses necessarily implies a two-step (sequential) estimation approach.

An alternate approach to specifying the form of $\sum$ is the latent class (finite mixture) version of Train's model proposed by Kamakura and Russell (1989). Utility parameters are estimated for each of several homogenous segments. The analysis is *a posteriori*; that is, the number of classes specified by the analyst is typically derived from the data analysis process and does not necessarily reflect *a priori* knowledge (Bollen, 2002).

A different approach to applying factors, referred to as latent variables, is taken by Ashok et al. (2002) and Walker (2001). Here the latent variables specifically represent characteristics of individuals, typically constructs like attitudes (e.g., satisfaction with past experiences). The latent variables are incorporated via a measurement model relating observed indicators (typically, rating scales), per conventional structural equations systems (e.g., Morikawa et al., 2002; Temme et al., 2008; Bolduc and Alvarez-Daziano, 2010; Yáñez et al., 2010; Hess and Stathopoulos, 2011). Proponents of this approach claim that it more closely captures choice processes by incorporating latent characteristics of decision makers. However, a statistical peculiarity of this approach is that the observed indicators of the latent characteristics are treated as endogenous and not used to make choice predictions.

SCM is consistent with random utility theory. It is a random coefficient model in which the coefficients for the covariates have a multivariate distribution with unknown parameters. It has factors that influence the random coefficients and can influence each other (i.e., structural equations link the factors). While the SCM approach can be more broadly applied, we typically describe it as a conditional logit model being applied to the attributes in a discrete choice experiment with

Gaussian random coefficients. It allows the specification of the form for $\sum$ and structure for the covariates in general, and the attributes in particular.

The approach and contribution is more easily demonstrated through a concrete example, as in Section 4. The example shows that SCM can be parsimonious and identified while testing *a priori* knowledge regarding individuals' choice behavior. However, the general model, and the more abstract mathematics, is presented first.

3. The model

Random utility theory specifies that for individual n alternative i has a utility U_{in} with a systematic component V_{in} and a idiosyncratic IID random error component ε_{in} where

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

On any one choice occasion the individual selects the alternative in the choice set with the highest utility. The individual can make different choices in situations which appear to an analyst to be identical. This apparent randomness is accounted for by the error term ε_{in}.

The probability of an alternative i being chosen from a choice set C, conditional on a choice being made, is known as the choice probability p. McFadden (1974, 2001) showed that under certain conditions for the error term, ε, the choice probability, p, is well approximated by the logit functional form, also known as a conditional logit model:

$$p_{in} = \frac{\exp(V_{in})}{\sum_{j \in C} \exp(V_{jn})} \tag{2}$$

The model we use for the systematic component of utility imposes no conditions on ε. Thus, other functional forms can be used. Conditional logit, as in Eq. (2), is used only for clarity and practical ease of estimation; it is not a necessary theoretical assumption of the model.

Traditionally, the systematic component of utility has been specified as a linear combination of covariates, X, as a row vector, being weighted by regression coefficients, β, as a column vector.

$$U_{in} = X_{in}\beta_n + \varepsilon_n \tag{3}$$

Walker (2001, pp. 177–178) has the general class of models with β as random and with factors, ξ, see Eq. (4). There are N individuals and the total number of alternatives considered by individual n is J_n:

$$U_n = X_n\beta_n + F_n\xi_n + v_n \tag{4}$$

where n is an individual, $n=1,\ldots, N$; i is an alternative, $i=1,\ldots, J_n$; U_n is a $(J_n \times 1)$ column vector of utilities; X_n is a $(J_n \times K)$ matrix of K covariates; β_n is a $(K \times 1)$ column vector of random components; ξ_n is a $(M \times 1)$ column vector of M factors; F_n is a $(J_n \times M)$ matrix of factor loadings and v_n is a $(J_n \times 1)$ column vector of idiosyncratic random utility errors. $\beta_n \sim N(\beta, \sum_\beta)$ and ξ_n is distributed as discussed below.

We modify the role and reduce the size of the factor matrix F_n by specifying that the factors act through the covariates:

$$U_n = X_n\beta_n + X_nF_n\xi_n + v_n \tag{5}$$

where F_n is now a smaller $(K \times M)$ matrix of factor loadings.

As a result the analyst has the ability to specify factors for the covariates, X, not just for the utilities, U. As discussed below, the analyst also can specify that any one factor acts only on a subset of the covariates and with different factors acting on different covariates. This is achieved through some factor loadings in F_n being specified by the analyst as unknown parameters and conversely some being specified to be zero. If required, factors can still act directly on the utilities for the alternatives, as in Eq. (4), by expanding the covariate matrix X to include columns of constants (ones). As a result Eq. (4) is a special case of Eq. (5).

Shifting from Eqs. (4) to (5), with the additional X_n, is a small change with deceptively large implications. In statistical terms the model no longer decomposes the variance of ε_n in Eq. (1), but instead decomposes and better specifies the variance of the coefficients for the covariates, X; leaving ε_n to capture the random utility error. In analytical terms, it is now possible to model associations (factors and structural equations) in the taste heterogeneity for specific attributes and levels. Implicit in Eq. (4) is that the alternatives can be named, in which case the attribute used in the naming can be included in the factors and structural equations. However, with Eq. (5) all attributes can be included equally and at the analyst's discretion. In behavioral terms the model with Eq. (5) recognizes that tastes vary not for alternatives but for attributes and with structure, not independently. With Eq. (5) this structure can be specified, allowing *a priori* knowledge to be tested.

Walker (2001, p. 180) adds a structural component:

$$\xi_n = \rho A_n \xi_n + T\zeta_n \tag{6}$$

where A_n is an $(M \times M)$ matrix of weights describing the influence of each factor, ξ, on the others. A_n can either be fixed or a function of unknown parameters, ρ is an unknown parameter, ζ_n is a $(M \times 1)$ column vector of random components with $\zeta_n \sim N(0,I)$ and T is a Cholesky matrix applying covariances to the random components.

This model contains the factor loadings in F_n and the structural equation parameters in A_n. We apply the constraint that both are not specific to the individual n and so lose the subscripts n. It is these unknown parameters that one wants to estimate for the population, not for individuals. Furthermore, because A is now constant, without loss of generalizability we specify $\rho=1$, although we leave to further research a discussion of the scale parameters (Fiebig et al., 2010). Thus, F_n is replaced with F and ρA_n is replaced with A. The model is

$$U_n = X_n\beta_n + X_nF\xi_n + v_n \tag{7}$$

$$\xi_n = A\xi_n + T\zeta_n \tag{8}$$

Solving Eq. (8) for ξ gives

$$\xi_n = (1-A)^{-1}T\zeta_n \tag{9}$$

and substituting in Eq. (7) gives

$$U_n = X_n\beta_n + X_nF(1-A)^{-1}T\zeta_n + v_n \tag{10}$$

Next we define the more general form of random coefficients, η_n, for the covariates as

$$\eta_n = \beta_n + F(1-A)^{-1}T\zeta_n \tag{11}$$

then, from Eq. (10):

$$U_n = X_n\eta_n + v_n \tag{12}$$

which is the structural choice model (SCM). The β_n in the model are traditionally referred to as the random coefficients, but we will (consistently) refer to them as regression coefficients. Our aim is to specify the form of $\sum$, which is the covariance matrix for the more general random coefficients, η_n. Finally, from Eq. (11) this covariance matrix, $\sum$, is

$$\Sigma = Cov(\eta_n) = Cov(\beta_n) + (F(1-A)^{-1}T)(F(1-A)^{-1}T)' \tag{13}$$

The parameters are summarized

Description	Symbol	Maximum number of unknown parameters
Factor loading	F	$K \times M$
Structural equation coefficient	A	$M \times M$
Constant: regression coefficient	$E[\beta_n]$	K
Random component: regression coefficient	$\mathrm{Cov}(\beta_n)$	$K \times (K+1)/2$
Random component: factor	$\mathrm{Cov}(T\zeta_n)$	$M \times (M+1)/2$

This general model is actually too general; hence, typically one would want to constrain it in some appropriate way. There are a large number of potential unknown parameters that cannot all be estimated from the data, including some that can be redundant, and not all are identified. In any one application each parameter is specified by the analyst to be either (i) unknown, and to be estimated from the data (if identified), or (ii) fixed to zero, or (iii) fixed to one. Hence the number of parameters to be estimated can be manageable and parsimonious. Also, there are only two sources of heterogeneity in the model, the regression coefficients β_n and the independent standard random components ζ_n. Both can be specified as random but in many applications they are not. In the example below in specification 1 neither is random, in 2 only β_n is random, in 3 only ζ_n is random, and in 4 both β_n and ζ_n are random. Furthermore, experience shows that due to the flexibility of SCM to create appropriate forms for $\sum$ and fit the data well, one often can constrain $\mathrm{Cov}(T\zeta_n)$ to be an identity matrix (i.e., the variances are one, there are no covariances, and T is an identity matrix).

It is now possible to specify the log likelihood function and estimate the parameters. From Eq. (10)

$$V_{in} = X_{in}\beta_n + X_{in}F(1-A)^{-1}T\zeta_n \tag{14}$$

Substituting in Eq. (2) gives the probability of a single choice. However, individual n makes a choice from each of several different choice sets. The additional information available through the joint probability of these choices (effectively, repeated measures) leads to the ability to identify the model. Let there be a choice sets C_1 to C_a each containing a finite number of alternatives. Let individual n select one alternative from each of these choice sets where the alternatives selected are c_1 to c_a. Then the joint probability of the a choices is

$$\Pr\left\{c_1,\ldots,c_a|C_1,\ldots,C_a\right\} = \iint \prod_{j=1}^{a} \frac{\exp(V_{c_jn}(\beta_n,\zeta_n))}{\sum_{i\in C_j}\exp(V_{in}(\beta_n,\zeta_n))} f(\beta_n)f(\zeta_n)d\beta_n d\zeta_n \tag{15}$$

The expression includes integrals that reflect the fact that the random components have known distributions but unknown values. Combining Eq. (15) over all individuals gives the likelihood function. The parameters can be estimated from the data using simulated maximum likelihood.

The model has K covariates, each with a random coefficient, η, where the covariance matrix for these coefficients, $\sum$, is of size $(K \times K)$. In the traditional random coefficient model $\sum$ can be problematic. It can be fully parameterized leading to $K \times (K+1)/2$ unknown parameters, which are too many and cannot be fully identified. $\sum$ can be diagonal, with K parameters, which is fewer, but is naïve, and still may not be fully identified. Instead, the structural choice model provides form such that $\sum$ can be fuller than diagonal while being identified and based on a manageable number of parameters. Next, we consider an example where structural choice models are used to test *a priori* knowledge of the nature of the taste heterogeneity for the attributes in a DCE.

4. Attracting Ph.D. students to jobs

The model is illustrated with data from a traditional DCE involving choices of employment options (job offers) made by a sample of PhD students at the 30 leading research universities in the United States.[3] The data were collected as part of a study conducted by the Harvard Graduate School of Education involving 797 doctoral participants. Seven attributes were varied (Table 1 describes attributes and associated levels). Each sample participant received eight choice sets containing two choice alternatives. Choice sets and choice alternatives were designed using the L^{MA} approach described in Louviere and Woodworth (1983) and Louviere et al. (2000). Participants made 4801 total choices.

A conditional logit model with fixed coefficients (in the model above the vector of parameters β has the same value for all individuals) was fitted to the data using effects coding for each attribute. The estimated partworths for each of the 26 attribute levels are given in Table 1 and the optimum log likelihood value was −2326. The range of partworths is a measure of the importance for each attribute; and the results suggest that the importance ordering is Geographic Location = 2.16, Rating of Department = 1.34, Salary = 1.12, Tenure = 1.04, Institution = 0.88, Contract = 0.82 and Balance of Work/life = 0.67.

The US university market is differentiated with some universities offering high Rating of Department and Institution and some high Geographic Location and only a few offering both. The research question we seek to address is whether all individuals are segmented on the same basis? That is, do individuals differ, tending to prefer either Departmental Rating or Geographic Location, but not both? And if so how do preferences for salary vary? Such questions about segmentation are routine in marketing and are often approached by observing continuous variables, such as attitude scales, and applying multivariate techniques such as factor analysis and structural equation modeling. Our objective is to apply the same style of analysis to model the taste heterogeneity in the DCE.

To emphasize the point, differentiation is a reflection of the alternatives available. It is a property of the choice set. Segmentation is a reflection of the patterns of consistency in individuals. It is a property of taste heterogeneity and $\sum$. Table 1 can be used in the traditional manner to evaluate the tradeoffs made between differentiated alternatives and attributes, but it cannot be used to explore dimensions of taste heterogeneity and the existence of segments. The latter require one to specify and estimate a mixed logit model and SCM.

The first step is to convert the data from effects coding to continuous coding. One covariate is created for each attribute, generating seven covariates in total. For each attribute the levels are scaled in proportion to the partworths in Table 1 varying from +1 for the highest partworth down to −1 for the lowest. The argument for this approach includes: (i) it is sensible when the levels of attributes are ordinal, as is the case in this DCE, (ii) it focuses the analysis on attribute taste heterogeneity and not on the separate levels of each attribute, (iii) the models are better identified (Walker, 2001, p. 54), and (iv) the output is more easily interpreted.

Four models are fit to the data. In specifications 1 (Fixed) and 2 (Random) utility is a function of β_n only; there are no factors. For specification 1 (Fixed) $\beta_n = \beta$ which is a vector of unknown constants to be estimated from the data and is equivalent to specifying $\beta_n \sim N(\beta, \sum_\beta = 0)$. In specification 2 (Random) there is a random component and $\beta_n \sim N(\beta, \sum_\beta)$ where $\sum_\beta$ is diagonal. The elements on the diagonal are unknown variances to be estimated from the data. The model and fit statistics are reported in Table 6. For the Fixed model the log likelihood and estimates of the parameters, in Table 2, are consistent with the earlier model applied to effects coding.

Of greater interest, in the analysis of taste heterogeneity, is the mixed logit model with random coefficients (in the model above the vector of parameters β has a unique value, β_n, for each individual n). The estimates of the parameters are:

In the Random model the vector β_n mediates the impact of the attributes. For larger values of $\beta_{n,j}$ attribute j will have greater impact on choice. Consider the first attribute, Balance of Work/life. Its covariate, X_1, takes on two values +1 and −1. If $\beta_{n,1} = 0.3$ then the contribution of the two levels of the attribute to utility is +0.3 and −0.3. However, if the coefficient is greater, say $\beta_{n,1} = 0.5$, then contribution increases, +0.5 and −0.5, and a change in the level of the attribute has a greater impact on choice. Over the population of individuals the vector β has Gaussian distributions that include negative values. An interesting example is the last attribute, Contract, where the standard deviation is large and many individuals will have a negative value for $\beta_{n,7}$. The interpretation is that shorter contracts are preferred by some individuals, where a possible explanation is that a long wait to convert to tenure is undesirable.

The Random model has high standard deviations, see Table 3, showing there is considerable taste heterogeneity, but the model cannot identify correlations between attributes and the multidimensional structure; the distributions of the vector β are specified to be independent and the covariance matrix $\sum$ is specified to be diagonal.

In contrast, specification 3 (Exploratory) looks for correlations (see Fig. 1). It applies an exploratory factor analysis with two correlated factors, ξ_1 and ξ_2; that is, $M = 2$. The factor matrix, F, of size $(K \times M) = (7 \times 2)$, contains 14 unknown factor loadings to be estimated from the data. The estimates in Table 4 provide a clear interpretation; the first factor relates to living 'conditions' and in particular Geographic Location while the second factor relates to 'rating', specifically, of the

[3] The authors thank the Harvard Graduate School of Education for the data used in the application.

Table 1
The experiment recorded doctoral participants' preferences for postdoctoral employment.

Attribute		Level		Partworth
1	Balance of work/life	1	Balance of work significantly different from your preference.	−0.33
		2	Balance of work matches your preference.	0.33
2	Tenure	1	Greater than 85% chance of tenure/contract renewal.	0.48
		2	71–85% chance of tenure/contract renewal.	0.15
		3	50–70% chance of tenure/contract renewal.	−0.08
		4	Less than 50% chance of tenure/contract renewal.	−0.56
3	Geographic Location	1	Somewhere you would really like to live.	0.95
		2	Somewhere you would be comfortable living.	0.53
		3	Somewhere you would be marginally satisfied with living.	−0.27
		4	Somewhere you would not like to live.	−1.21
4	Rating of Department	1	Department rated among the top 10 in your discipline.	0.62
		2	Department rated between 11 and 20 in your discipline.	0.22
		3	Department rated between 21 and 40 in your discipline.	−0.11
		4	Department not in top 40 in your discipline.	−0.72
5	Rating of Institution	1	Institution rated among the top 10 in the U.S.	0.28
		2	Institution rated between 10 and 20 in the U.S.	0.14
		3	Institution rated between 21 and 40 in the U.S.	−0.01
		4	Institution not in top 40 in the U.S.	−0.41
6	Salary	1	At least 25% above average for your discipline.	0.50
		2	At least 10% above average for your discipline.	0.08
		3	About average for your discipline.	0.04
		4	At least 15% below average for your discipline.	−0.62
7	Contract	1	10 years.	0.37
		2	5 years.	0.12
		3	3 years.	−0.05
		4	1 year.	−0.44

Table 2
Parameter estimates for the Fixed model with continuous coding.

Attribute	**Estimate of β**	**se**
Balance of work/life	0.33	0.022
Tenure	0.52	0.038
Geographic Location	1.08	0.041
Rating of Department	0.67	0.038
Rating of Institution	0.35	0.036
Salary	0.56	0.037
Contract	0.41	0.037

Table 3
Parameter estimates for the Random model with continuous coding.

Attribute	**Estimate of mean for β**	**se**	**Estimate of std. dev. for β**	**se**
Balance of work/life	0.60	0.05	0.33	0.085
Tenure	0.92	0.09	0.50	0.163
Geographic Location	2.07	0.15	1.44	0.140
Rating of Department	1.10	0.09	0.95	0.126
Rating of Institution	0.57	0.07	0.58	0.135
Salary	1.01	0.09	0.89	0.133
Contract	0.61	0.08	1.03	0.129

Department. However, the estimate of the correlation between the two factors is zero (and not significant). Finally, Salary, which earlier was identified as important, is common to both factors. Before interpreting these results a new model, specification 4 (Confirmatory), is fitted to the data.

Specification 4 (Confirmatory) captures the role of Salary, see Fig. 2. It divides the attributes into two mutually exclusive confirmatory factors for ‘conditions’ and ‘rating’ and then analyses the impact of each on a third factor, ‘salary’; thus, $M=3$. The factor matrix, F, of size $(K\times M)=(7\times 3)$, contains 6 unknown parameters to be estimated from the data. One parameter

Systematic component of utility for individual n:

$$\begin{array}{rll} V_n & = X_{n,1}(\beta_1 + f_{1,1}\xi_{n,1} + f_{1,2}\xi_{n,2}) & \text{Work/life} \\ & + X_{n,2}(\beta_2 + f_{2,1}\xi_{n,1} + f_{2,2}\xi_{n,2}) & \text{Tenure} \\ & + X_{n,3}(\beta_3 + f_{3,1}\xi_{n,1} + f_{3,2}\xi_{n,2}) & \text{Geographic} \\ & + X_{n,4}(\beta_4 + f_{4,1}\xi_{n,1} + f_{4,2}\xi_{n,2}) & \text{Department} \\ & + X_{n,5}(\beta_5 + f_{5,1}\xi_{n,1} + f_{5,2}\xi_{n,2}) & \text{Institute} \\ & + X_{n,6}(\beta_6 + f_{6,1}\xi_{n,1} + f_{6,2}\xi_{n,2}) & \text{Salary} \\ & + X_{n,7}(\beta_7 + f_{7,1}\xi_{n,1} + f_{7,2}\xi_{n,2}) & \text{Contract} \end{array}$$

V_n is a column vector containing all alternatives relevant to individual n.
$X_{n,1}$ to $X_{n,7}$ are the 7 columns of the matrix of covariates X_n.
The 7 coefficientsβ_1 to β_7 areunknown and have the same values for all individuals, n.
There are 14 unknown factor loadings f.
There are two factors $\xi_{n,1}$ and $\xi_{n,2}$ where

$$\begin{aligned} \xi_{n,1} &= \zeta_{n,1} \\ \xi_{n,2} &= \zeta_{n,2} \end{aligned}$$

$\zeta_{n,1}$ and $\zeta_{n,2}$ are random components with Gaussian distribution, means zero and standard deviations one and correlation φ.

Fig. 1. Specification 3: Exploratory factor analysis with 2 factors.

Table 4
The factor loadings (the f parameters) in the Explanatory model show that factor 1 reflects 'conditions' and factor 2 'rating', but both reflect salary.

Attribute	Factor 1 'Conditions'	Factor 2 'Rating'
Salary	0.4	0.3
Geographic Location	1.1	−0.2
Balance of work/life	0.2	0.0
Tenure	0.2	0.0
Contract	−0.3	−0.1
Rating of Department	0.0	0.7
Rating of Institution	0.0	0.1

is fixed to a value of one (otherwise the model cannot be fully identified), and the remaining 14 are specified to be zero.

$$F = \begin{pmatrix} f_{1,1} & 0 & 0 \\ f_{2,1} & 0 & 0 \\ f_{3,1} & 0 & 0 \\ 0 & f_{4,2} & 0 \\ 0 & f_{5,2} & 0 \\ 0 & 0 & 1 \\ f_{7,1} & 0 & 0 \end{pmatrix}$$

The structural equation matrix A, of size $(M \times M)=(3 \times 3)$, contains 2 unknown parameters to be estimated from the data and the remaining 7 are specified to be zero:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ a_{3,1} & a_{3,2} & 0 \end{pmatrix}$$

Estimates of the parameters are in Table 5 and Fig. 3. The results confirm that individuals can be segmented on the basis of the two dimensions, 'conditions' and 'rating', where 'conditions' is dominated by Geographic Location and 'rating' by Rating of Department. Again the two factors are not correlated. Those preferring 'conditions' are no less, or more, likely to

Systematic component of utility for individual n:

$V_n =$	$X_{n,1}\left(\beta_{n,1} + f_{1,1}\xi_{n,1}\right)$	Work/life
$+$	$X_{n,2}\left(\beta_{n,2} + f_{2,1}\xi_{n,1}\right)$	Tenure
$+$	$X_{n,3}\left(\beta_{n,3} + f_{3,1}\xi_{n,1}\right)$	Geographic
$+$	$X_{n,4}\left(\beta_{n,4} + f_{4,2}\xi_{n,2}\right)$	Department
$+$	$X_{n,5}\left(\beta_{n,5} + f_{5,2}\xi_{n,2}\right)$	Institute
$+$	$X_{n,6}\left(\beta_{n,6} + \xi_{n,3}\right)$	Salary
$+$	$X_{n,7}\left(\beta_{n,7} + f_{7,1}\xi_{n,1}\right)$	Contract

V_n is a column vector containing all alternatives relevant to individual n.
$X_{n,1}$ to $X_{n,7}$ are the 7 columns of the matrix of covariates X_n.
The 7 coefficients $\beta_{n,1}$ to $\beta_{n,7}$ are are unknown and constant for each individual n but vary over individuals.
There are 6 unknown factor loadings f.
There are three factors $\xi_{n,1}$, $\xi_{n,2}$ and $\xi_{n,2}$ where

$\xi_{n,1} =$	$\zeta_{n,1}$	'Conditions'
$\xi_{n,2} =$	$\zeta_{n,2}$	'Rating'
$\xi_{n,3} =$	$a_{3,1}\xi_{n,1} + a_{3,2}\xi_{n,2}$	'Salary'

There are 2 unknown structural coefficients a.
$\zeta_{n,1}$ and $\zeta_{n,2}$ are random components with Gaussian distribution, means zero and standard deviations one and correlation φ.

Fig. 2. Specification 4: Confirmatory model.

Table 5
Parameter estimates for the Confirmatory model.

	Attribute		Estimate of mean for β	se	Estimate of std. dev. for β	se
β_1	Work/life		0.64	0.06	0.28	0.09
β_2	Tenure		0.99	0.10	0.53	0.15
β_3	Geography		2.17	0.16	0.00	0.50
β_4	Department		1.14	0.10	0.00	0.80
β_5	Institute		0.59	0.07	0.59	0.14
β_6	Salary		1.12	0.10	0.68	0.16
β_7	Contract		0.61	0.09	1.00	0.13
	From factor to	**Attribute**	**Estimate of *f***	**se**		
$f_{1,1}$	'Conditions'	Work/life	0.21	0.05		
$f_{2,1}$	'Conditions'	Tenure	0.19	0.10		
$f_{3,1}$	'Conditions'	Geography	1.51	0.15		
$f_{7,1}$	'Conditions'	Contract	−0.15	0.12		
$f_{4,2}$	'Rating'	Department	1.03	0.13		
$f_{5,2}$	'Rating'	Institute	0.07	0.10		
$f_{6,3}$	'Salary'	Salary	Fixed to 1			
	From factor to	**Factor**	**Estimate of *a***	**se**		
$a_{3,1}$	'Conditions'	'Salary'	0.47	0.12		
$a_{3,2}$	'Rating'	'Salary'	0.46	0.14		
	From factor to	**Factor**	**Estimate of φ**	**se**		
φ	'Conditions'	'Rating'	−0.10	0.22		

prefer 'rating'. Both dimensions link to 'salary' equally (both have an equal concern for salary). The model generates *R*-squared statistics. Of the taste heterogeneity in the attribute Salary 46% is explained by the two factors 'conditions' and 'rating'.

The analysis started with the observation that the market might be seen as being differentiated along the lines of Rating of Department and Institution versus Geographic Location; *a priori* we proposed that segmentation might be on the same basis. The analysis shows that Rating of Institution is quite unimportant whereas Rating of Department is important as is Geographic Location. The analysis shows there is considerable taste heterogeneity for all the attributes but the two principal factors in the multidimensional structure are based on Rating of Department and Geographic Location. Preferences for these two attributes are independent. The market is segmented on preference for Rating of Department and preference for Geographic Location but it is not accurate to conclude that individuals prefer either one or the other and not both. Furthermore, Salary is equally important to those who prefer Rating of Department and those who prefer Geographic Location.

SCMs were used above to conduct an exploratory factor analysis, a confirmatory factor analysis and to fit structural equations involving all the attributes. Models of this nature are well known in SEM but are new to the analysis of taste heterogeneity in DCEs. They are prevalent in the analysis of continuous variables but not to the specific type of data generated by DCEs where discrete choices are made from choice sets that differ. The new outcomes from SCM for DCEs include a parsimonious and full specification for the covariance matrix, $\sum$ (see Table 10).The SCM approach also introduces a new role for DCEs in testing competing *a priori* knowledge and hypotheses on the structure of taste heterogeneity. In the DCE, Fig. 3 confirms *a priori* knowledge regarding segmentation on the basis of conditions and rating but disconfirms that individuals necessarily favor only one or the other. Had prior knowledge suggested other competing structures, the SCM analysis also could have tested the competing models. This approach in operationalizing and testing theory for the structure of taste heterogeneity is new to DCEs

5. Diagnostics

SCM provides an array of diagnostics (Rungie, 2011). From the AIC and Likelihood ratio tests in Table 6 the Confirmatory model is preferred, although the Random model performs well on the BIC measure. The means for the random coefficients η

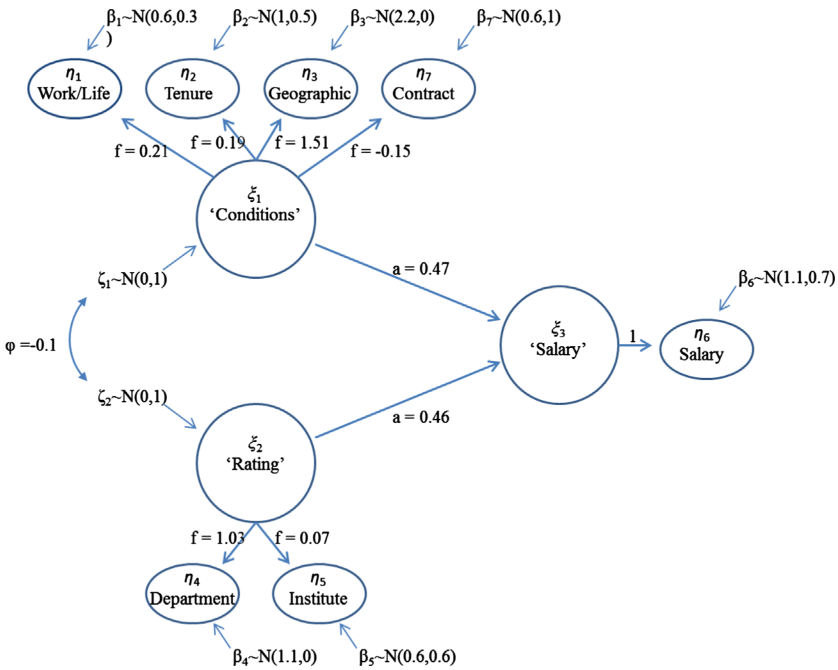

Fig. 3. Path diagram for the Confirmatory model.

Table 6
Harvard case study: the five competing specifications of the model.

Specification	1 Fixed	2 Random	3 Exploratory	4 Confirmatory
Number of covariates, $K=$	7	7	7	7
Number of factors, $M=$			2	3
F			14	6
A				2
$E[\beta_n]$	7	7	7	7
$\mathrm{Cov}(\beta_n)$		7		7
$\mathrm{Cov}(T\zeta_n)$			1 a	1 a
Total	7	14	22	23
Log likelihood	−2326	−2212	−2214	−2185
AIC	4665	4452	4472	4417
BIC	4698	4517	4575	4525
LR Test c.f. specification 1		$p=2E-45$	$p=4E-39$	$p=3E-50$
LR Test c.f. specification 2				$p=3E-08$

Note: There is one correlation and no other unknown parameters: ζ_n are independent standard Gaussian and T provides the one correlation.

for each covariate X are restated in Table 7 and the variances are in Table 8. While proofs are yet to be developed, we suspect that choice models are biased to smaller variances. When fully identified models estimate higher variances they are likely to be better models that capture more of the taste heterogeneity in the data. On the basis of the results in Tables 6 and 8 the Confirmatory model is preferred.

In particular, the analysis indicates that the Exploratory model is not a good fit to the data. For completeness its correlation matrix is given in Table 9 and unsurprisingly it reports high correlations because this type of model is designed to emphasize correlations.

The correlation matrix for the Confirmatory model is given in Table 10. Consistent with the earlier findings it shows Geographic Location and Rating of Department both correlate with Salary but not with each other.

For SCM, as with SEM (Bollen, 1989), effective identification is established by examining the properties of models and following a series of guidelines (Rungie et al., 2011). First, each model presented above has face validity. In the Confirmatory model there is the potential for redundant parameters. As a result some parameters rather than being free, to be estimated

Table 7
Means of the random coefficients η for the four models.

	Fixed	Random	Exploratory	Confirmatory
Balance of work/life	0.33	0.60	0.46	0.12
Tenure	0.52	0.92	0.73	0.32
Geographic Location	1.08	2.07	1.63	2.28
Rating of Department	0.67	1.10	0.88	1.06
Rating of Institution	0.35	0.57	0.46	0.35
Salary	0.56	1.01	0.83	0.86
Contract	0.41	0.61	0.40	1.02

Table 8
Variance of the random coefficients η for the four models.

	Fixed	Random	Exploratory	Confirmatory
Balance of work/life	0	0.11	0.03	0.12
Tenure	0	0.25	0.04	0.32
Geographic Location	0	2.09	1.29	2.28
Rating of Department	0	0.91	0.52	1.06
Rating of Institution	0	0.34	0.00	0.35
Salary	0	0.80	0.26	0.86
Contract	0	1.06	0.08	1.02

Table 9
Exploratory model: correlation matrix (extracted from $\sum$) for the random coefficients η.

	Work/life	Tenure	Geography	Department	Institute	Salary	Contract
Work/life							
Tenure	0.95						
Geography	1.00	0.96					
Department	−0.17	0.14	−0.14				
Institute	−0.21	0.10	−0.18	1.00			
Salary	0.55	0.78	0.58	0.73	0.70		
Contract	−0.90	-0.99	−0.91	−0.28	−0.23	−0.86	

Table 10
Confirmatory model: correlation matrix (extracted from $\sum$) for the random coefficients η.

	Work/life	Tenure	Geography	Department	Institute	Salary	Contract
Work/life							
Tenure	0.20						
Geography	0.59	0.34					
Department	−0.06	−0.03	−0.10				
Institute	−0.01	0.00	−0.01	0.11			
Salary	0.27	0.16	0.46	0.44	0.05		
Contract	−0.09	−0.05	−0.15	0.01	0.00	−0.07	

from the data, instead are fixed. For 'salary' the factor loading $f_{6,2}$ is fixed to 1 and the random components ζ_3 is omitted (fixed to zero). Secondly, for each model the Hessian matrix is well-behaved; the inverse exists and has positive values in the leading diagonal. Thirdly, standard errors throughout are small (a poorly identified parameter will have a large standard error). Exceptions include estimates of the standard deviations of β for Geographic Location and Rating of Department in Table 10. The results suggest that these two parameters are effectively identified but they make little contribution to the fit of the model. Finally, the estimation was well-behaved: conversion was quick and all parameter estimates are sensible.

6. Conclusion

In the DCE the *a priori* structure of the attributes has separate forms. Firstly, differences between individuals in their involvement and engagement in the subject and domain of the DCE, postdoctoral employment issues, will lead to differences in how much they differentiate globally between all levels of all attributes. Secondly, the seven attributes reflect a smaller set of over-arching meta-attributes that will be reflected in correlations between attributes and in segments. However, *a priori* knowledge suggested a particular structure for this correlation that was first explored through specification 3 and then confirmed through specification 4. The analysis confirmed part of the prior knowledge and disconfirmed part. The DCE demonstrates how *a priori* knowledge about the structure of the taste heterogeneity for attributes can be operationalized, and in turn lead to specifications that fit the data better and have greater validity.

The structural choice models can be parsimonious. The number of unknown parameters to be estimated from the data in the four specifications is relatively small, particularly when compared to extending the classic random coefficient model by directly parameterizing the correlations in $\sum$. Specifications 1–4 are all identified. The inclusion of the factors in specifications 3 and 4 and structural equations in 4 did not compromise identification.

The four specifications are competing models. They demonstrate how structural choice models can test and evaluate competing *a priori* knowledge. They give an indication of potential new contributions to the development of the theory in the subject area and domain of DCEs and to the theory of choice.

The structural choice model is a general model with one functional form. All four specifications for the DCE are just that, different specifications of the one model. This property has two useful implications. First, the need for a special software program is avoided. The four specifications were fit to the data by providing different input matrices to a single computer program. Potentially all analysts can share and use the same programs. Second, the model has a standard language and notation that can assist communication between analysts when working on the same DCE and even on different DCEs.

The structural choice model is a random coefficient model where the covariance matrix, $\sum$, has form. The model can be parsimonious, valid, identified, and fit the data better. It can reflect *a priori* knowledge on the structure of the attributes. It can test, evaluate, and develop the theory for the subject area and domain of DCEs. The model follows the call by Adamowicz et al. (2008) to develop new behavioral insights and theory about choice and choice processes. There are likely many more applications.

References

Adamowicz, W., Bunch, D., Cameron, T.A., Bellaert, B.G.C., Manneman, M., Keane, M., Louviere, J., Meyer, R., Steenburgh, T., Swait, J., 2008. Behavioral frontiers in choice modeling. Marketing Letters 19, 215–228.

Ashok, K., Dillon, W.R., Yuan, S., 2002. Extending discrete choice models to incorporate attitudinal and other latent variables. Journal of Marketing Research 39 (1), 31–46.

Ben-Akiva, M., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, MA.

Ben-Akiva, M., Morikawa, T., 1990. Estimation of switching models from revealed preferences and state intentions. Transportation Research A 24A (6), 485–495.

Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., Polydoropoulou, A., 2002. Integration of choice and latent variable models. In: Mahmassani, H.S. (Ed.), In Perpetual Motion: Travel Behavior Research Opportunities and Application Challenges, Elsevier Science, Oxford, UK, pp. 431–470.

Bolduc, D., Alvarez-Daziano, R., 2010. On estimation of hybrid choice models. In: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice. Emerald Group Publishing Limited, Bingley, UK.

Bollen, K.A., 1989. Structural Equations with Latent Variables. John Wiley & Sons, New York.

Bollen, K.A., 2002. Latent variables in psychology and the social sciences. Annual Review of Psychology 53, 605–634.

Dunson, D.B., 2000. Bayesian latent variable models for clustered mixed outcomes. Journal of the Royal Statistical Society B 62 (2), 355–366.

Elrod, T., Keane, M.P., 1995. A factor-analytic probit model for representing the market structure in panel data. Journal of Marketing Research 32 (1), 1–16.

Fiebig, D.G., Keane, M.P., Louviere, J., Wasi, N., 2010. A generalized multinomial logit model: accounting for scale and coefficient heterogeneity. Marketing Science 29 (3 (May–June)), 393–421.

Hensher, D.A., Louviere, J., Swait, J., 1999. Combining sources of preference data. Journal of Econometrics 89 (1–2), 197–221.

Hensher, D.A., Rose, J.M., Greene, W.H., 2005. Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge, UK.

Hess, S., Stathopoulos, A., 2011. Linking Response Quality to Survey Engagement: A Combined Random Scale and Latent Variable Approach. Institute for Transport Studies, University of Leeds, Leads.

Jöreskog, K.G., 1970. A general method for analysis of covariance structures. Biometrika 57, 239–251.

Jöreskog, K.G., 1973. A general method for estimating a linear structural equation system. In: Goldberger, A.S., Duncan, O.D. (Eds.), Structural Equation Models in the Social Sciences, Seminar Press, New York, pp. 85–112.

Jöreskog, K.G., Sörbom, D., 1996. LISREL 8: User's Reference Guide. Scientific Software International, Lincolnwood, IL.

Kamakura, W.A., Russell, G.J., 1989. A probabilistic choice model for market segmentation and elasticity structure. Journal of Marketing Research XXVI, 379–390.

Keane, M.P., 1997. Modeling heterogeneity and state dependence in consumer choice behavior. Journal of Business & Economic Statistics 15 (3), 310–327.

Lee, S.Y., 2007. Structural Equation Modeling: A Bayesian Approach. John Wiley & Sons Ltd., West Sussex, England.

Louviere, J.J., Woodworth, G., 1983. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. Journal of Marketing Research 20, 350–367.
Louviere, J.J., Hensher, D.A., Swait, J., 1999. Conjoint analysis methods in the broader context of preference elicitation methods. In: Gustafson, A., Hermann, A., Huber, F. (Eds.), Conjoint Measurement: Methods and Applications. Springer, Berlin, pp. 279–318.
Louviere, J.J., Hensher, D.A., Swait Jr., J., 2000. Stated Choice Methods. Cambridge University Press, New York.
Louviere, J.J., Street, D., Carson, R., Ainslie, A., DeShazo, J.R., Cameron, T., Hensher, D.A., Kohn, R., Marley, A.A.J., 2002. Dissecting the random component of utility. Marketing Letters 13 (3), 177–193.
Luce, R.D., 1959. Individual Choice Behavior: A Theoretical Analysis. Wiley, New York.
Marschak, J., 1960. Binary choice constraints on random utility indicator. In: Arrow, K.I., Karlin, S., Suppes, P. (Eds.), Stanford Symposium on Mathematical Methods in the Social Sciences, Stanford University Press, Standford, CA, pp. 312–329.
Maydeu-Olivares, A., Böckenholt, O., 2005. Structural equation modeling of paired-comparison and ranking data. Psychological methods 10, 285–304.
McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics, Academic Press, New York, pp. 105–142.
McFadden, D., 2001. Economic choices. American Economic Review 91, 351–378.
McFadden, D., Train, K.E., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15, 447–470.
Morikawa, T., Ben-Akiva, M., McFadden, D., 2002. Discrete choice models incorporating revealed preference and psychometric data. Econometric Models in Marketing 16, 29–55.
Rungie, C., 2011. Structural Choice Modelling Software Manual DisCoS. AdelaideUniversity of South Australia.
Rungie, C., Coote, L.V., Louviere, J.J., 2011. Structural choice modelling: theory and applications to combining choice experiments. Journal of Choice Modelling 4 (3), 1–29.
Temme, D., Paulssen, M., Dannewald, T., 2008. Incorporating latent variables into discrete choice models: a simultaneous estimation approach using SEM software. Business Research 1 (2), 220–237.
Thurstone, L.L., 1927. A law of comparative judgment. Psychological Review 34, 273–286.
Train, K.E., 2003. Discrete Choice Models with Simulation. Cambridge University Press, New York.
Walker, J.L., 2001. Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables. Unpublished Doctoral Dissertation. Department of Civil & Environmental Engineering, Massachusetts Institute of Technology.
Yáñez, M., Raveau, S., de Dios Ortúzar, J., 2010. Inclusion of latent variables in mixed logit models: modelling and forecasting. Transportation Research Part A: Policy and Practice 44 (9), 744–753.