UNIVERSITY OF WAIKATO

Hamilton

New Zealand

A Monte Carlo Evaluation of the Logit-Mixed Logit under Asymmetry and Multimodality

Cristiano Franceschinis, Riccardo Scarpa and Mara Thiene

Working Paper in Economics 23/17

October 2017

Corresponding Author

Riccardo Scarpa

School of Accounting, Finance
and Economics
University of Waikato
Private Bag 3105
Hamilton
NEW ZEALAND, 3240

Email: riccardo.scarpa@waikato.ac.nz

Cristiano Franceschinis

Department of Land
and Agro-Forest Environments
University of Padova
Padova
ITALY

Mara Thiene

Department of Land
and Agro-Forest Environments
University of Padova
Padova
ITALY

Abstract

The recently introduced (Train 2016) logit-mixed logit (LML) model is a key advancement in choice modelling: it generalizes many previous parametric and semi-nonparametric methods to represent taste heterogeneity for bundled nonmarket goods and services. We report results from Monte Carlo experiments designed to assess performance across workable sample sizes and to retrieve data-driven random coefficients distributions in the three variants of the LML model proposed in the seminal paper. Assuming a multi-modal data generating process, with a panel of four and eight choices per respondent, we compare the performance of WTP-space LML models with conventional parametric model specifications based on the Mixed logit model with normals (MXL-N) in preference and WTP space. Results are encouraging and support the adoption of flexible LML specifications with a high number of parameters as they seem to do better, but only at large enough sample sizes. To explore the saliency of the Monte Carlo results in an empirical application, we use data obtained from a discrete choice experiment to derive preferences for tap water quality in the province of Vicenza (northern Italy). LML models retrieve multimodal and asymmetric distributions of marginal WTPs for water quality attributes. Results show not only how the shape of such distributions vary across tap water attributes, but also the importance of being able to uncover them, considering that they would be hidden when using the MNL-N.

Keywords

logit-mixed logit
flexible taste distributions
panel random utility models

Introduction

Preference heterogeneity affects welfare estimates obtained from discrete choice models. The effects are particularly salient when such models are based on surveys developed for nonmarket valuation studies. A case in point is their use to inform the process of definition and negotiation of service tariffs in regulated industries. Tap water supply is a complex quasi-public good jointly managed by water utilities and regulatory bodies (Willis and Scarpa, 2002; Willis et al., 2005; Hensher et al., 2005; Scarpa et al., 2007; Rungie et al., 2014; Thiene et al., 2015) as natural monopolies. Regulatory bodies periodically require water utilities to evaluate the adequacy of their tariffs to the economics of long term investment plans. Thus, water utilities and their regulatory authorities are interested in gathering information about customer preference in order to strategically define investment in infrastructure for water delivery, water treatment and sewer services. If a water factor service produces benefits to costumers of water utilities, this is deemed worth investing on. But knowing the distribution features of the benefits is also important. A high estimate of mean benefit might justify a comparatively higher investment in securing the regularity of service in that specific water factor. But knowing if the high benefit mean is underpinned by a bimodal distribution with a mode at a relatively lower benefit value and a perhaps smaller mode at a higher level of benefits, or instead by a single modal value centred on the mean, might induce different forms of strategic investments. An understanding of the distributional features of customer preference is therefore important.

Since the seminal work by Train (1998) the choice modelling approach has proved to be an insightful way to investigate consumer preferences from choice survey data. In addition to its popularity amongst transport and health analysts, this approach has been embraced by applied economists who used it in a wide range of nonmarket valuations, such as in food quality and safety (Balcombe and Fraser, 2011; Campbell and Doherty, 2012); landscape (Scarpa et al., 2009; de Ayala et al., 2015; Nordén et al., 2017); water services (Hanley et al., 2006; Rigby et al., 2010; Brouwer et al., 2015; Thiene et al., 2015); ecosystems services (Ohdoko and Yoshida, 2012; Thiene et al., 2012; Johnston et al., 2013; Chaikaew et al., 2017); recreational activities (e.g. Thiene and Scarpa, 2009; Jacobsen and Thorsen, 2010; Juutinen et al., 2014; Mejía and Brandt, 2015; Bertram et al., 2017; Morey and Thiene, 2017); and energy resources (e.g. Scarpa and Willis, 2010; Willis et al., 2011; Yoo and Ready, 2014; Yamamoto, 2015; Boeri and Longo, 2017; Bartczak et al., 2017).

Over the past few decades, the field of choice modelling has witnessed a substantial development in model specifications accounting for various forms of taste heterogeneity. Taste variations can be decomposed into observed and unobserved preference heterogeneity, depending on whether such heterogeneity can be explained by observable factors or not. Analysts, who model decision-making processes using random utility maximization theory, cannot include all relevant factors that determine differences in taste across people. Inevitably only part of the differences in taste are 'observable'–in the sense that their variation is associated to measurable variables–while the remainder is due to unobservable or unmeasurable factors.

Before the advent of mixed logit models, the workhorse of this type of literature used to be the Multinomial Logit (MNL) model McFadden (1973). This can capture observed preference heterogeneity typically by interacting attributes with characteristics of the individual, but treatment of unobserved preference heterogeneity requires further assumptions and a model extension. To take into account unobserved preference heterogeneity, Boyd and Mellman (1980) introduced the mixed Mixed Logit (MXL) model by taking the MNL specification and adding the assumption of the presence of random parameters that follow a pre-specified parametric, continuous mixing distribution. MXL became the new standard practice in choice modeling after McFadden and Train (2000) showed that any random utility maximization process can be approximated up to an arbitrary level by a MXL model, if mixing distributions of random parameters are specified correctly. A contributory factor was the growth in the computational power of microcomputers, that made the execution of estimation via simulation viable to most.

One issue that has received attention from the beginning is how to adequately characterize taste distributions (Wedel et al., 1999; Hensher and Greene, 2003). A quick perusal of recent publications can show that many applications still use parametric distributions (either bounded or unbounded), in which at least some of the distributional features (e.g. part of their shape) are restricted *a priori* by the analyst's choice of distribution. For example, unbounded distributions by construction may have long tails, significant density around zero values and may allow parameters to take either positive or negative values, even when there are theoretical expectations on their signs. This causes several problems: coefficient estimates with high density on value intervals with theoretically unwarranted signs or with implausibly extreme values (too large or too small to be realistic) may be obtained, which can jeopardize the credibility of the results. For example, our literature review of the top 5 journals in environmental economics shows that in the period 2012-2017 as many as 83 papers use mixed logit with assumptions of normal distributions for the random parameters.

Bounded distributions overcome some of these problems (Train and Sonnier, 2005), but still assume a well-shaped distribution and can overestimate the true mean toward the bounds, biasing the welfare measures (Cherchi and Polak, 2005). Furthermore, when the range of the cost coefficient has any positive density in proximity of zero, boundedness of distributions for the utility coefficients in preference space settings, may not imply boundedness of the implied WTP, which can still assume unrealistic values (Scarpa, Thiene and Train, 2008). The most recent and promising advances adopt semi-parametric or nonparametric models to capture the randomness in individuals' tastes. Here, the shape of the distribution is unknown and defined as part of the estimation process. Further flexibility is necessary and this can be provided by semi-parametric and non-parametric approaches, which are typically data-hungrier.

Several approaches have been proposed in the literature to estimate semiparametric or nonparametric random coefficients in discrete choice models. Abe (1999) introduced a framework to estimate semiparametric utility functions within the Multinomial Logit model (MNL). With this methodology, splines can be used to model non-linear influences on the response variable. While this model requires no a priori assumption about the functional form, the amount of smoothness for each function (in terms of the equivalent degrees of freedom) must be fixed before estimation. This approach has been extended in Kneib et al. (2007), in which estimation of smoothing parameters is based on a mixed model representation of penalized splines. Fukuda and Yai (2011), similarly, proposed a semi-nonparametric model based on smoothing splines, by specifying cubic spline functions for each explanatory variable.

Li (2011) proposed a semi-parametric choice model based on the B-splines approach developed in Eilers and Marx (1996). Fosgerau and Bierlaire (2007), instead, proposed a method to approximate any continuous distribution using a Legendre polynomial. The use of polynomials is a very flexible method to retrieve preference heterogeneity because different distributions can be recovered simply by adding more terms to the series expansion. Fosgerau and Hess (2007) compared the semi-nonparametric Legendre-polynomial logit with other parametric logit models and found the semi-nonparametric specification best in terms of retrieving the true distribution of the random parameters (with true distributions ranging from Uniform to multi-modal). Such observation is expected because of the flexibility of specifying a higher number of parameters in semi-nonparametric approaches.

Scarpa, Thiene and Marangon (2008) used the semi-nonparametric polynomial approach and found that the use of a flexible taste distribution increased the plausibility of the retrieved form of taste heterogeneity in their data, which emerged as bimodal, with modes at both sides of zero, rather than unimodal centred on zero. The implication being that rather than indifference to the attribute there was a population split between those who deemed it mildly desirable and those who—instead—deemed it undesirable. Bajari et al. (2007) proposed a method that takes advantage of a linear regression-type specification. The authors assume that the population can be sorted into finite classes or clusters (i.e. discrete number of preference parameters) and assert that their estimator is non-parametric because any mixing distribution can be approximated by making the number of classes large enough. However, this linear regression method may violate some necessary constraints on the model parameters. To handle this issue, Fox et al. (2011) re-parameterized MXL and derived a specification similar to that of Bajari et al. (2007), but used inequality constrained linear least squares.

Train (2008) used computation-efficient EM algorithms for non-parametric estimation of random parameter logit-type models. Fosgerau and Mabit (2013) suggest drawing random numbers from some initial distribution (e.g. uniform) and transform these draws using a polynomial or other function to generate the mixing distribution. Similarly, Bastin et al. (2010) proposed a non-parametric method to approximate the inverse cumulative distribution function of the mixing distribution. They use a polynomial approximation of an initially chosen uniform distribution. A major limitation of these procedures is the need of understanding the relationship between the shape of the mixing distribution and that of the initial distributions.

Train (2016) has recently proposed the semi-nonparametric logit-mixed logit (LML) model. As the name suggests, this model contains two logit formulations: one for the decision maker's probability to choose an alternative and another for the probability of selecting a parameter from a finite parameter space. The exponential terms in the latter logit formulation ensure a positive probability and the denominator ensures normalization, i.e. all probabilities sum to one. In addition, the shape of the logarithm of the mixing distribution can be defined by different type of functions such as polynomials, step functions, and splines), among many others. This estimator has been supplied with general purpose code in MatLab and presents very desirable computational features, further examined and confirmed in Bansal et al. (2016).

The objective of this paper is twofold. Since LML provides a generalized and flexible framework for semi-nonparametric mixed logit models, the number of observations and of parameters required in order to adequately retrieve the underlying features of random taste heterogeneity are worth exploring. Thus, in the first part of this paper, we conduct a Monte Carlo study to answer this questions. Specifically, we compare results from LML models to those retrieved from traditional parametric specifications, such as the MXL with normal distributions (MXL-N). The ability of LML to retrieve parametric distributions (normal, log-normal, uniform, symmetric bimodal normal, uniform, discrete and discrete log-normal) in specifications with utility in preference-space has already been studied elsewhere (Bansal et al., 2016). We extend this work by focussing on DGP with asymmetric bimodality and trimodality and on utility in WTP-space because of their importance in applied welfare analysis for public goods and their relation to median voter behaviour, and hence political markets for public good provision (e.g. see the discussion in Mitchell and Carson, 1989). In the context of LML, we also investigate the finding of Fosgerau and Hess (2007), which suggests that the ability to recover an underlying distribution depends on the number of parameters in the mixing distribution, i.e. a higher number of parameters yields a better approximation of the true distribution.

In addition to the Monte Carlo studies, we provide an empirical case study in which standard parametric approaches lead to overlooking some features that instead emerge as important with LML. Specifically, we analyze the preferences of householders in a part of the province of Vicenza (North Italy) for tap water attributes. The objective of this empirical application is to explore the implications of alternative LML specifications with varying number of parameters on the estimates of the distributions of WTP values for the improvement of tap water services. Since the true distribution of WTP is not known, we compare the distributions of WTP estimates of LML with MXL and explicitly state the benefits of using LML over parametric specifications.

The remaining paper is organized as follows: section 2 illustrates MXL and LML models, section 3 describes the Monte Carlo experiment design, Section 4 discusses simulation results, section 5 illustrates the empirical study and its results. Finally, section 6 draws the conclusions of the paper.

Econometric modeling

The Mixed Logit Model (MXL) with normals

The MXL model represents random taste heterogeneity by allowing for different preference parameters for each decision-maker (Revelt and Train, 1998). The utility derived by individual n from choosing alternative j in choice occasion t is logit:

$$U_{njt} = \mathbf{x}_{njt}{}'\boldsymbol{\beta}_n + \varepsilon_{njt}, \text{ where } n = 1, ..., N; J = 1, ..., J; t = 1, ..., T, \tag{1}$$

and where $\boldsymbol{\beta}_n$ is a vector of parameters for individual n which is assumed to follow a continuous mixing distribution in the population; $\mathbf{x}_{njt}$ is a conformable column vector of observed attributes of alternative i; ε_{njt} is the error term assumed to follow a Gumbel distribution. The conditional probability $P_n(jt|\boldsymbol{\beta}_n)$ of individual n choosing alternative j in choice occasion t is:

$$P_n(jt|\boldsymbol{\beta}_n) = \frac{\exp(\mathbf{x}_{njt}{}'\boldsymbol{\beta}_n)}{\sum_{k=1}^{J} \exp(\mathbf{x}_{nkt}{}'\boldsymbol{\beta}_n)}. \tag{2}$$

Many variants of the MXL models can be obtained by assuming different mixing distributions of the random parameters. The most common is the MXL-N that imposes a multivariate normal mixing distribution, i.e., $\boldsymbol{\beta}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $y_{nkt} = 1$ if individual i chooses alternative j in choice situation t, and 0 otherwise. For a panel of T choices, the unconditional probability of the sequence of T preferred alternatives by individual n is facing J alternatives is:

$$P_n(jT|\boldsymbol{\beta},\boldsymbol{\Sigma}) = \int \left\{ \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(\mathbf{x}_{njt}'\boldsymbol{\beta}_n)}{\sum_{k=1}^{J} \exp(\mathbf{x}_{nkt}'\boldsymbol{\beta}_n)} \right]^{y_{njt}} \right\} f(\boldsymbol{\beta}_n|\boldsymbol{\mu},\boldsymbol{\Sigma}) d\boldsymbol{\beta}_n, \tag{3}$$

where $f(\boldsymbol{\beta}_n|\boldsymbol{\mu},\boldsymbol{\Sigma})$ is the probability density function with mean hyperparameter vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$ for the random taste parameters $\boldsymbol{\beta}_n$. Hyperparameters in the MXL-N model are estimated through the maximum simulated likelihood estimator (Gouríeroux and Monfont, 1996).

The Logit-Mixed Logit Model (LML)

In LML models Train (2016), the joint mixing distribution of the random parameters $\boldsymbol{\beta}_n$ is assumed to be discrete over a finite support set S. Discretization is not a constraint because the support set is essentially a multidimensional grid that can be made larger and denser by considering a broader domain of parameters and a higher number of grid points. The joint probability mass function of random parameters in LML is specified by the following logit formula:

$$w_n(\boldsymbol{\beta}_r|\boldsymbol{\alpha}) = Pr(\boldsymbol{\beta}_n = \boldsymbol{\beta}_r) = \frac{\exp(\mathbf{z}(\boldsymbol{\beta}_r)'\boldsymbol{\alpha})}{\sum_{s\in S} \exp(\mathbf{z}(\boldsymbol{\beta}_s)'\boldsymbol{\alpha})}, \tag{4}$$

where $\boldsymbol{\alpha}$ is a vector of parameters, $z(\boldsymbol{\beta}_r)$ defines the shape of the mixing distribution, and r denotes the point in the grid for the evaluation of $\boldsymbol{\beta}$. The unconditional probability of the sequence of choices of individual n is the following weighted sum:

$$P_n(jT|\boldsymbol{\alpha}) = \sum_{r\in S} \left\{ \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(\mathbf{x}_{njt}'\boldsymbol{\beta}_r)}{\sum_{k=1}^{J} \exp(\mathbf{x}_{nkt}'\boldsymbol{\beta}_r)} \right]^{y_{njt}} \right\} w_n(\boldsymbol{\beta}_r|\boldsymbol{\alpha}). \tag{5}$$

In LML models, the vector $\boldsymbol{\alpha}$ is estimated using the maximum likelihood estimation procedure. Inclusion of all the points of the support set in the estimation of LML is unnecessary and computationally expensive, so a subset of points is drawn within S. The logit formula in Eq. 4 to compute probability mass of random parameters results into an efficient computation of the gradient of the sample loglikelihood, facilitating gradient-based methods in estimation.

The z functions in LML models

A critical issue in LML model is the specification of the S variables that describe the mixing distribution. Following Train (2016), we adopt three different functions: i) polynomials (LML-Poly), ii) step function (LML-Step function), iii) spline (LML-Spline).

An important feature of LML-Poly is that many of the standard distributions can be approximated by varying the order of the polynomial. For example, Train (2016) shows that the normal distribution can be introduced in LML framework by considering $z(\boldsymbol{\beta}_r)$ to be a second order polynomial of a special form. The polynomial can be extended to higher orders to gain greater flexibility of the mixing distribution, bearing in mind that the number of inflection points is equal to the polynomial order minus one. Among the various categories of polynomials, orthogonal polynomials (e.g. Legendre, Hermite, Jacobi, Chebyshev, Bernstein etc.) have the advantage of having uncorrelated terms. Dependence among the elements of multi-dimensional $\boldsymbol{\beta}$ can still be captured by cross-products of the terms of each element's polynomial.

A second alternative to define $z(\boldsymbol{\beta}_r)$ is represented by a step function based on a grid over the parameter ranges (i.e. the support set S). Suppose S is partitioned into M subsets, labelled as T_m where $m = \{1, 2, ..., M\}$. Let the probability mass function $W(\boldsymbol{\beta})$ be the same for all points within each subset, but different among subsets. The logit formula for the probability masses in then:

$$w_n(\boldsymbol{\beta}_r|\boldsymbol{\alpha}) = Pr(\boldsymbol{\beta}_n = \boldsymbol{\beta}_r) = \frac{\exp(\sum_{m=1}^{M} \alpha_m I(\boldsymbol{\beta}_r \in T_m))}{\sum_{s\in S} \exp(\sum_{m=1}^{M} \alpha_m I(\boldsymbol{\beta}_s \in T_m))}. \tag{6}$$

The z variables are the M indicators which identify the subset containing $\boldsymbol{\beta}_r$. If the subsets do not overlap, then one of the coefficients is normalized to zero. With overlapping subsets, instead, one coefficient is normalized to zero for each possible way of covering the set S. In LML-Step function the number of estimated parameters is equal to the number of grid points.

Finally, a linear spline can be used to define $\mathbf{z}(\beta)$, once defined over h knots. Spline functions connect piece-wise polynomial functions at a high degree of smoothness and in a linear setting they can be written in the form $\boldsymbol{\alpha}'\mathbf{z}(\beta)$, as needed in the LML specification. Consider a simple example of spline with $h = 2$ and with starting point at β_1, ending point in β_4, and place the two knots at β_2 and β_3, with $\beta_1 < \beta_3 < \beta_3 < \beta_4$. Let the corresponding elements of the vector $\boldsymbol{\alpha}$ define the spline heights. The elements of vector $\mathbf{z}(\beta)$ in this case are:

$$\begin{cases} z_1(\beta) = \left(1 - \frac{\beta - \bar{\beta}_1}{\bar{\beta}_2 - \bar{\beta}_1}\right) I(\beta \le \bar{\beta}_2) \\ z_2(\beta) = \left(\frac{\beta - \bar{\beta}_1}{\bar{\beta}_2 - \bar{\beta}_1}\right) I(\beta \le \bar{\beta}_2) + \left(1 - \frac{\beta - \bar{\beta}_2}{\bar{\beta}_3 - \bar{\beta}_2}\right) I(\bar{\beta}_2 < \beta \le \bar{\beta}_3) \\ z_3(\beta) = \left(\frac{\beta - \bar{\beta}_2}{\bar{\beta}_3 - \bar{\beta}_2}\right) I(\bar{\beta}_2 < \beta \le \bar{\beta}_3) + \left(1 - \frac{\beta - \bar{\beta}_3}{\bar{\beta}_4 - \bar{\beta}_3}\right) I(\beta_3 < \beta) \\ z_4(\beta) = \left(\frac{\beta - \bar{\beta}_3}{\bar{\beta}_4 - \bar{\beta}_3}\right) I(\beta_3 < \beta) \end{cases}, \tag{7}$$

where $I(\cdot)$ is an indicator function. In LML-Spline, the number of parameters is equal to the number of knots plus one.

Monte Carlo experiment design

To assess the performances of different model specifications, we conducted a Monte Carlo study based on three attributes. The first and the second attribute are assumed to be non-monetary, whereas the third is assumed to be the price attribute. The two non-monetary attributes were coded as dummy variables, taking the values of 0 and 1, indicating presence or absence in the alternative they describe. The price attribute was assumed to have two levels as well, having the values of 1 and 2. The Monte Carlo experiment was developed under the assumption that the true data generation processes (DGPs) are asymmetric and bi- and trimodal, and utility specified in WTP space, so that coefficients are interpretable as marginal WTPs. Consistently with random utility theory, it was assumed that a respondent chooses the alternative with maximum utility between the two alternatives. The utility of respondent n for alternative j in choice occasion t was specified as:

$$U_{njt}(\beta_n) = \lambda_n^*(\omega_n^1 x_{njt}^1 + \omega_n^2 x_{njt}^2 - p_{njt}) + \varepsilon_{njt}, \tag{8}$$

where λ_n^* is the price/scale coefficient and ω_n^1 and ω_n^2 are the marginal WTPs for attribute 1 and attribute 2. To compare performance of MXL-N and LML models at increasing level of complexity of $mWTP$ distributions, we generated two data generating processes (DGP). In the first set, DGP 1, ω_n^1 and ω_n^2 were assumed to follow a bimodal distribution, obtained by mixing two normal distributions, whereas the price/scale coefficient λ_n^* was assumed to follow a mixture of two log-normal distributions (Figure 1 upper panel), to ensure its positive sign. The price coefficient p_i was assumed to be fixed to -1. Error term ε_{njt} was assumed to follow a standard Gumbel distribution. The distribution parameters in DGP 1 were specified as asymmetric bimodal, as follows:

$$\omega_n^1 \sim \mathcal{N}(\boldsymbol{\mu}^1, \boldsymbol{\Sigma}^1) \text{ with } \boldsymbol{\mu}^1 = \begin{bmatrix} 0.5 \\ 1.2 \end{bmatrix} \boldsymbol{\Sigma}^1 = \begin{bmatrix} 0.04 & 0 \\ 0 & 0.04 \end{bmatrix} \text{ with } \Pr = \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}, \tag{9}$$

$$\omega_n^2 \sim \mathcal{N}(\boldsymbol{\mu}^2, \boldsymbol{\Sigma}^2) \text{ with } \boldsymbol{\mu}^2 = \begin{bmatrix} -1.5 \\ 1.5 \end{bmatrix} \boldsymbol{\Sigma}^2 = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.25 \end{bmatrix} \text{ with } \Pr = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \tag{10}$$

$$\lambda_n^* = \exp(\theta), \theta \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \text{ with } \boldsymbol{\mu} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \boldsymbol{\Sigma} = \begin{bmatrix} 0.25 & 0 \\ 0 & 1.0 \end{bmatrix} \text{ with } P = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}. \tag{11}$$

The shape of the distributions for both random coefficients (or $mWTPs$) of attributes in DGP 1, ω^1 and ω^2, are shown in the upper panel of Figure 1. Note that in the distribution for ω^1 the two modal densities differ: that for low benefits is much higher than that for higher benefits. Also note that in the distribution for ω^2 one mode is negative and has higher density than the positive mode, to denote asymmetric distributions of winners and losers linked to the supply of that binary level attribute. It is intuitive to conclude that these two forms of asymmetric bimodality in the population distribution of benefits from a public good provision will lead to importantly differences in policy actions if they were confounded with a unimodal distribution with some intermediate modal value.

In the DGP 2, we assumed $mWTP$ have an asymmetric trimodal distribution, obtained as a mixture of three normals (reported in Figure 1, lower panel). The distribution parameters were given the following values:

$$\omega_n^1 \sim \mathcal{N}(\boldsymbol{\mu}^1, \boldsymbol{\Sigma}^1) \text{ with } \boldsymbol{\mu}^1 = \begin{bmatrix} 1.5 \\ 3.5 \\ -1.2 \end{bmatrix} \boldsymbol{\Sigma}^1 = \begin{bmatrix} 0.04 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0.40 \end{bmatrix} \text{ with } P = \begin{bmatrix} 0.3 \\ 0.3 \\ 0.4 \end{bmatrix}, \tag{12}$$

$$\omega_n^2 \sim \mathcal{N}(\boldsymbol{\mu}^2, \boldsymbol{\Sigma}^2) \text{ with } \boldsymbol{\mu}^2 = \begin{bmatrix} 5.5 \\ 3.0 \\ 1.2 \end{bmatrix} \boldsymbol{\Sigma}^2 = \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0.09 \end{bmatrix} \text{ with } P = \begin{bmatrix} 0.3 \\ 0.3 \\ 0.4 \end{bmatrix}, \quad (13)$$

$$\lambda_n^* = \exp(\theta), \theta \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \text{ with } \boldsymbol{\mu} = \begin{bmatrix} 0.1 \\ 1.2 \end{bmatrix} \boldsymbol{\Sigma} = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.02 \end{bmatrix} \text{ with } P = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}. \quad (14)$$

The shape of the distributions for both random $mWTP$s of attributes in DGP 2, ω^1 and ω^2, are shown in the lower panel of Figure 1. This time, in the distribution for ω^1 the two positive modal densities differ less than in DGP 1 and are contrasted by a highest density around a negative mode. The distribution for ω^2 has only positive modal values, but with three different densities, the highest of which is at low level of benefits, accompanied by two similar level densities at higher benefit levels. As for DGP 1, it is intuitive to conclude that these two forms of asymmetric trimodality in the benefits distribution for a public good provision will also lead to vastly different optimal policy actions.

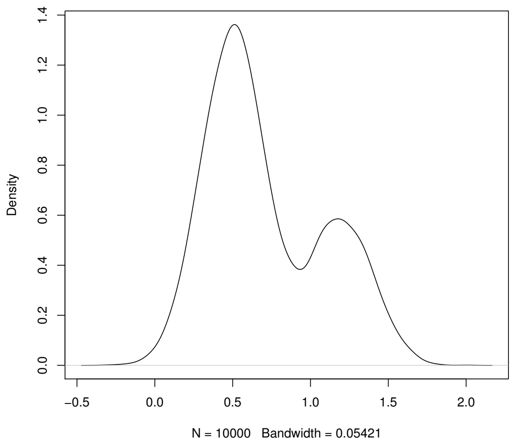
For each of the $r = 1, \ldots, 1,000$ simulated sets of discrete choice responses, we estimated 14 models. These models consist of one MXL-N in preference space with normal distributions for each non-price attribute, one MXL-N in WTP space with normal coefficients for all non-price attributes, four LML-Poly with varying number of parameters (12, 24, 36, 48), four LML-Step with varying number of steps (12, 24, 36, 48), and four LML-Spline with varying number of knots (12, 24, 36, 48). All LML models were in WTP-space and all price coefficients were log-normal. Data generation and all estimations were performed in MatLab using Train's code adjusted to our purpose.

In estimation choice probabilities were simulated in the sample log-likelihood with 250 Halton draws in all models. To simulate the sampling distributions properties of $mWTP$ values from the MXL in preference space, 10,000 draws were taken from the estimated distribution of each non-monetary attribute coefficient. Each draw was then divided by a draw from the estimated distribution of the cost coefficient. Standard statistics for the distribution of these WTPs were then calculated for these draws (but see the caveats in Daly et al. (2012)). Bearing in mind that the mean squared error (MSE) of an unbiased estimator equals its variance, to compare the performance of different models we computed the MSE over the 1,000 experiments, as well as the mean relative absolute error (MRAE) of the estimates, computing:

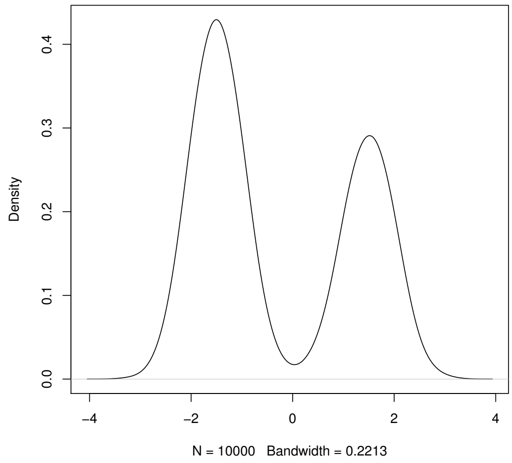
$$MSE = \frac{1}{R}\sum_{r=1}^{R}(\hat{\omega}_r - \omega)^2, r = 1, \ldots, 1,000 \quad (15)$$

$$MRAE = \frac{1}{R}\sum_{r=1}^{R}\left|\frac{\hat{\omega}_r - \omega}{\omega}\right|, r = 1, \ldots, 1,000 \quad (16)$$

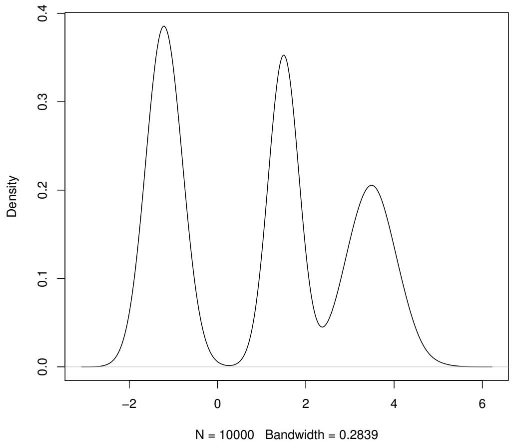
where ω is the WTP value used in the data generating process and $\hat{\omega}_r$ is the value estimated from the r^{th} Monte Carlo experiment. Furthermore, the local maxima and minima of coefficients' distributions retrieved in each experiment were computed and compared, in order to assess the capability of different model specifications to retrieve good approximations of the underlying true distribution from DGP 1 and DGP 2. In



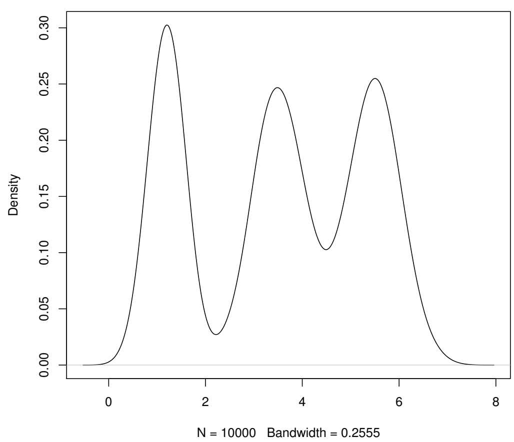
ω^1 in DGP 1 (bimodal)



ω^2 in DGP 1 (bimodal)



ω^1 in DGP 2 (trimodal)



ω^2 in DGP 2 (trimodal)

Figure 1: Kernel smoothing plots of $mWTP$ in the 2 DGPs.

both DGP 1 and DGP 2 we used three experimental designs: i) D-error minimizing design, ii) random design, iii) full factorial design. Note that under our asymmetric multimodal DGPs the MXL-N is obviously biased, but it might nevertheless tradeoff bias with lower variance under some conditions.

Results from the Monte Carlo experiment

The first important observation concerns the results obtained from different experimental designs. We found that results are consistent across designs: the relative performances of the various specifications are very similar. So, to save space, in what follows we describe the results with focus on those obtained from datasets generated with the D-error minimizing design. Similar results were also obtained with regards to the accuracy measures MSE and $MRAE$ and for this reason in the remainder we limit our discussion to the MSE values across different models and DGPs. All omitted results are available from the authors upon request.

Short panel results, $T = 4$, bimodal distributions

Table 1 reports MSE values for estimates of the mean $mWTP$ for attribute 1 and attribute 2 as retrieved from datasets with four choice tasks per respondent, with DGP 1 that implemented asymmetric bimodal distributions of the real parameters. It is immediately noticeable that, given ω, the value of MSE decreases as N increases: accuracy is increased by larger samples. For small samples (simulated respondents $N = 70$ and $N = 210$) the best performing model—that is, the one with the lowest MSE (and $MRAE$)—is the MXL-N WTP space, which outperforms all LML models. A bias variance tradeoff seems to take place at this level. At intermediate sample sizes ($N = 490$ and $N = 980$ simulated respondents) some of the LML specifications outperformed the MXL in WTP-space (e.g. LML-poly with $\kappa = 36$ for ω^1 and LML-Spline with $\kappa = 48$ and $\kappa = 24$ for ω_2), but it is only at large sample sizes ($N = 1,960$) that LML models consistently outperformed the MNL-N WTP for some value of κ. At $N = 1,960$ there is also a clear improvement in performance by LML models with higher dimensions of κ. Among LML models based on step functions and splines the best model specifications were those with $\kappa = 48$, whereas the best model specification among LML-Poly models was the one with $\kappa = 36$ according to both MSE and $MRAE$.

In terms of identification of the optimal number of parameters κ to be adopted in LML models for both bimodal coefficients, we obtain no clear indication at such sample sizes. According to the MSE values for ω^1, for the LML-Poly the best specification is the one with $\kappa = 24$, followed by $\kappa = 48$ and then $\kappa = 36$. Moving to the results for LML-Step, the best performing models are those with high number of κ (36 and 48). Finally, among LML-Spline, the best performing model specification is the one with $\kappa = 24$, followed by the one with $\kappa = 36$. For to the MSE for the second coefficient ω_2, the best performing LML-Poly has $\kappa = 24$ and 48; for the LML-Step $\kappa = 48$, while for the LML-Spline is the one with $\kappa = 36$.

The second important distribution feature is its spread, often measured by the standard deviation. The MSE for these statistics of the Monte Carlo results are reported in Table 2. As for the means, at the smallest sample size the MXL-N WTP outperforms all models (and it always outperforms the MXL-N in preference space), but already at $N = 210$ we have LML-Step with $\kappa = 36$ that does better and at higher sample sizes LML models do better both more frequently and more consistently, especially at high values of κ.

Long panel results, $T = 8$, bimodal distributions

Tables 3 and 4 reports the same statistics as above, but for the longer panel with eight choice tasks per respondent ($T = 8$). So, the number of choices are doubled at each sample size. Doubling the number of responses collected from each respondent obviously sharpens the estimation of the distributions of taste, as it allows for both better panel designs and more information from more numerous choices. Whether and at what sample size this difference becomes apparent with respect to $T = 4$ is an empirical question we try to answer here. The results from datasets with small sample size ($N = 70$ and $N = 210$) are similar to those retrieved from datasets with four choice scenarios per respondent, in that the MXL model outperforms the LML specifications and there are no clear indications about the effect of increasing the number of parameters of LML specifications.

However, *MSE* for both means and standard deviations show that flexible LML specifications consistently surpass the MXL model at both intermediate and large sample sizes. Similarly to the short panel results, for $N = 980$ respondents, each LML specification outperformed the MXL model for some value of a κ. This suggests that increasing the number of observations per respondent (a longer panel) does not seem to allow analysts to retrieve substantively more accurate estimates with LML models at smaller sample sizes. At both $N = 980$ and $N = 1,960$, it is also apparent that model specifications with large κ outperform the others.

Short panel results, $T = 4$, trimodal distributions

We now move to the results for the choice data generated under the DGP 2 with asymmetric trimodal distributions for ω_n^1 and ω_n^2 reported in Tables 5 and 6 for the case with short panel. Results are similar to those retrieved for the first set of coefficients in that the MXL-N WTP model always outperforms the MNL-N Pref. and does so for LML models at small sample sizes. The main difference is that in this case, already a $N = 490$, so at intermediate sample sizes, the *MSE* for LML are frequently smaller than those for the MXL-N WTP. It seems to be the case that with a trimodal distribution DGP flexible distribution models are more accurate than MXL-N at lower sample sizes, even with short panel, especially the LML-Step and LML-Spline.

Long panel results, $T = 8$, trimodal distributions

Tables 7 and 8 report the same statistics for the long panel. No noticeable difference is found from the results obtained for the short panel, indicating that doubling the number of choices per respondent does not substantially change the tradeoff between bias and sample size.

Bimodality

Tables 9-12 report the means and standard deviations of modal estimates of distributions of ω_n^1 and ω_n^2 from the various model specifications in both the short panels and long panels. They all have in common the bimodal DGP 1 as true process.

The first important observation concerns the number of modal values retrieved from different model specification. Naturally, MXL-N models (both in preference and WTP space) are inherently unimodal and cannot, by their very nature, imply bimodal distributions, but they are expected to retrieve a mean/mode/median at an intermediate value between the modes of the underlying DGPs. Indeed the results confirm this. Instead, LML models can retrieve bimodal distributions and do so in our experiment, with a degree of accuracy that increases with the sample size. This confirms that LML models are able to approximate better the shape of the true underlying distributions of random coefficients, and should always be considered when unimodality is not well supported a-priori, as it is often the case.

The second objective of the analysis was to identify how close the local maxima and minima retrieved from different LML specification were to the true ones. In this sense, it appears that increasing both the sample size and κ increases the accuracy of the estimates. In fact, the values that are closer to the real ones were obtained from LML specifications with $\kappa = 48$ estimated using datasets with $N = 1,980$. Of course, one can also compute MSE and $MRAE$ values for modal estimates and compare them across LML models. We have those results, but chose not to discuss them here.

Trimodality

Tables Tables 13-16 report the number of modal values from model estimated on data from DGP 2 (trimodal real distributions of mWTPs). As for the bimodal case, MXL model cannot retrieve the complex form of the real distributions, and deliver unimodal distributions at intermediate values of the modes in the real data. LML specifications with $\kappa = 12$, instead, always retrieve distributions with two modal values, instead of three. On the other hand, LML specifications with $\kappa = 48$ always correctly retrieve distribution with three modal values. Finally, LML specifications with intermediate $\kappa = 24-36$ retrieve distributions with three modal values at intermediate and large sample sizes, but bimodal distributions at lower N. As in previous cases, it is apparent that increasing sample sizes and κ increases the accuracy of estimates. In fact, modal values of distributions retrieved from model estimated from large datasets are closer to the DGP values.

Overall the results suggest that LML models may outperform the standard MXL-N specifications and represent more accurately complex distributions, but do so especially at large N. With regards to the optimal κ to be used in LML models, it seems that high κ values should be considered, but unsurprisingly they work better at large N.

Empirical application

To add saliency to the Monte Carlo results,we applied the estimator to an empirical application based on a discrete choice experiment (DCE) focused on household preferences for tap water attributes in the province of Vicenza (northern Italy).

The area under investigation is knwon as a tannery district. In fact, it is the most important district of that type in Italy and one of the most important in Europe. It accounts for nearly one third of fine European leather production (UNCI, 2010). The leather industry is a potential big polluter, due to the fact that a large amount of water is required to treat hides which are preserved using salts and usually travel from South America. Consequently, water emissions from hides treatment plant affect freshwater quality in the area. Historically this industry was located at the foothills of the Alps and it prospered here because of the several artesian springs providing a regular flow of one of the most pristine water sources in Italy. Water pollutants are present in low concentrations in hides, but may have high toxicity as tanning processes make use of toxic heavy metals like chrome and other chemical pollutants (e.g., sulphate and sodium chloride). The current charging system for public wastewater processing is based on threshold concentrations of contaminants per unit of volume of water used, rather than on total discharged load of contaminants. Hence, large amounts of pristine water from local springs are used to dilute concentrations of industrial pollutants. To give a sense of proportion, the capacity of the local sewage plant is sufficient for a population of 1.5 million, while the local population is only about 115,000 residents. Thus, information about householders' preferences for tap water attributes is crucial for local authorities in order to strategically set water tariffs and plan investments. Much of the necessary treatment infrastructure would otherwise benefit tanneries, which would then be heavily subsidized by residential water users, causing a major misallocation of resources.

The DCE was based on five attributes, namely:

i) the frequency with which chlorine odor can be smelled in water use (daily, once a week, once a month, never or always),

ii) the frequency with which chlorine taste could be tasted in the water (same frequencies as for odor),

iii) turbidity due to fine air bubbles (absent, low, medium or high turbidity),

iv) calcium carbonate staining in pipes (presence/absence of staining), and

v) the cost attribute, which was described as the additional yearly amount of money that a household would pay (in water bills) at current consumption levels.

The experimental design adopted in the study was based on the criterion of Bayesian D-error minimization (Sándor and Wedel, 2001; Ferrini and Scarpa, 2007; Rose and Bliemer, 2009) where the error was computed at parameter estimates obtained from a preliminary prior study of 80 households based on a design orthogonal on the differences. The point estimates from the pilot study informed the prior distribution for the Bayesian design, and the standard errors defined the variances of the prior distributions, which were assumed normal. Probabilities were derived from a simulation based on 200 Halton draws, and used to construct a final design using Ngene (ChoiceMetrics, 2009). The designed resulted in 36 choice tasks, blocked into four orthogonal blocks of nine choice tasks each.

Using the datasets obtained with the CE, we estimated 16 model specifications. These models consist of:

- one MXL-N in preference space,
- one MXL-N in WTP space,
- two latent class models with respectively two and three classes, to capture perfectly correlated multimodality,
- four LML-Poly with varying dimensions of κ (22, 33, 44, 55),
- four LML-Step with varying dimensions of κ (22, 33, 44, 55),
- and four LML-Spline with varying dimensions of κ (22, 33, 44, 55).

Models were estimated using MatLab code available from K. Train website and choice probabilities were simulated in the sample log-likelihood with 250 Halton draws. To compare performance across models with different number of parameters we report the simulated log-likelihood at convergence ($\ln \mathcal{L}^*$), the Akaike information criteria (AIC) and the Bayesian information criteria (BIC) are reported. Given the importance of multimodality in this context, we also report the number of modal values of the estimated distributions of random coefficients ($mWTP$).

Given the Monte Carlo results and the large number of observations in our dataset, we expect the LML models to outperform the MXL-N ones in terms of fit to the data. We also expect that LML specifications with large number of parameters to outperform those with small number of parameters. Finally, we expect LML specifications (especially those with large number of parameters) to be able to retrieve the features of real underlying distributions, even when these are quite complex, such as asymmetric and multimodal.

Model fit and estimated modes

Table 17 reports the model fit statistics for all models. All the information criteria favor LML specifications, as compared to the MXL and LC specifications. The results also support the previous finding of an increase of model performance at large κ at this sample size. In terms of performance across different z functions within LML, the LML-Spline specification emerges as the best performing when based on $\kappa = 55$, according to the AIC, but when based on $\kappa = 44$ according to the BIC, which applies a heavier penalty on over-parameterization. A close second in fit is the LML-Step, which is also best at $\kappa = 55$, according to the AIC, but at $\kappa = 44$ according to the BIC. In third position we find LML-Poly, and in this case both AIC and BIC converge in indicating $\kappa = 55$ a the model with best fit.

For the sake of space we only report and discuss the multimodality aspect of the results. Table 18 reports the estimated modes of $mWTP$ distributions. Obviously, MXL-N retrieved unimodal distributions in all random coefficients. LML models with $\kappa = 22$ and $\kappa = 33$, instead, retrieved bimodal distributions for most of the coefficients. In particular, LML-Poly with $\kappa = 22$ retrieved bimodal distributions for seven coefficients and with $\kappa = 33$ did so for eight $mWTP$ distributions. LML-Step $\kappa = 22$ retrieved bimodal distributions for seven parameters and LML-Step with $\kappa = 33$ for nine $mWTP$ distributions. Similar number of bimodality are found in the estimates from LML-spline.

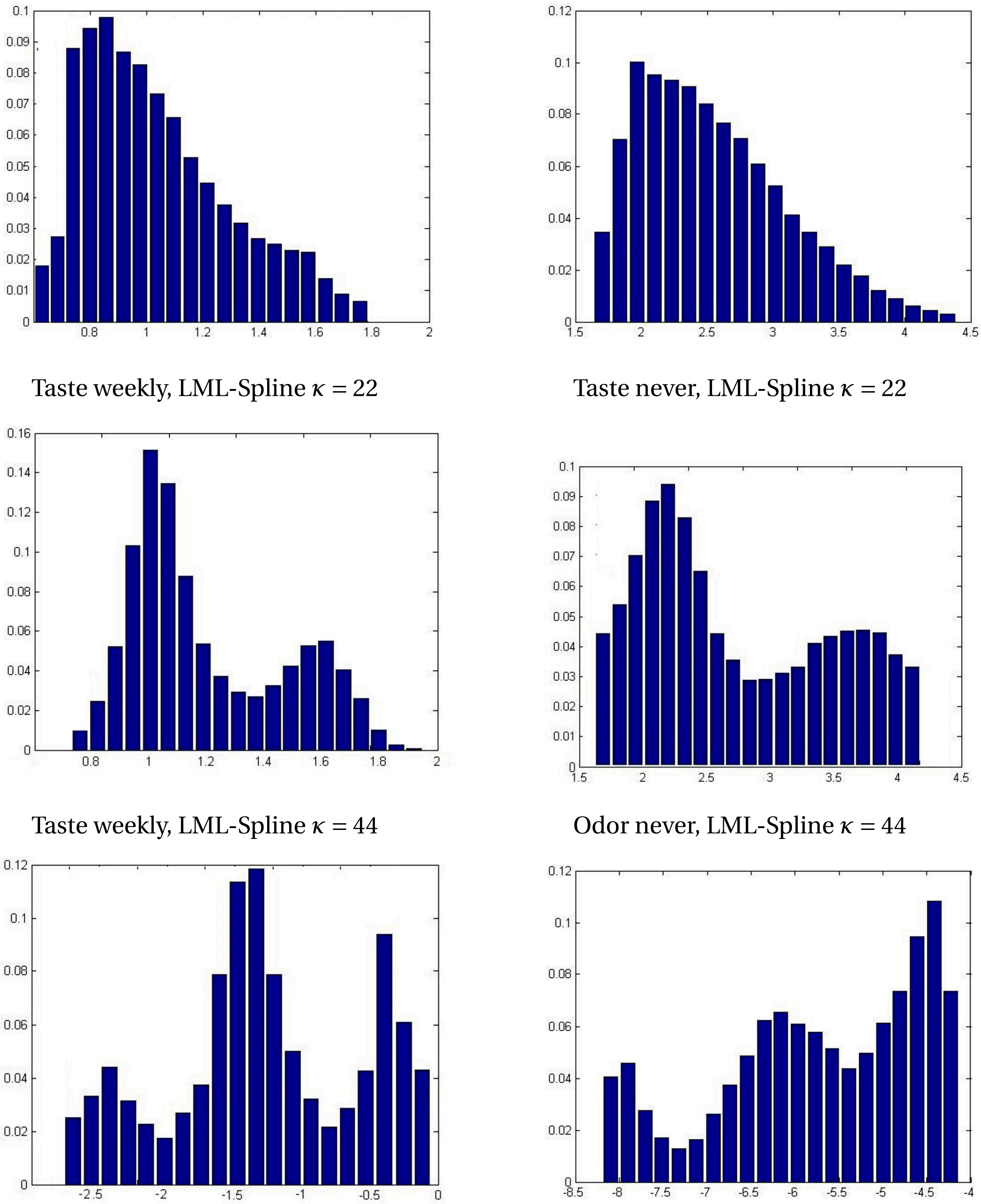
The histograms reported in the first and second rows of figure 2 are a good illustration of the effect of an increase in κ on the estimated multimodality of the random WTP for taste with weekly frequency and never. While with $\kappa = 22$ the two attributes appear to have unimodal distributions, with $\kappa = 44$ they appear bimodal.

Distributions with three modal values where retrieved only from LML models with $\kappa = 44$ and $\kappa = 55$ (e.g. see the bottom histograms in figure 2 for mild and extreme turbidity). In particular, all the specifications with such number of parameters retrieved tri-modal distributions for chlorine odor once per month, chlorine taste once per month, medium and extra degrees of turbidity. All this information would be lost in MXL-N, and plausibly in all conventional parametric distributions. We note that some multimodality can be captured in means of individual-specific distributions, but those statistics are of difficult interpretation at the population level (see chapter 11 in Train, 2009, for a discussion).

Conclusions

This paper provides results from a Monte Carlo experiment and an empirical application conducted to investigate the ability of different variants of the recently proposed Logit-mixed Logit (LML) in retrieving the underlying heterogeneity distributions of random parameters, with a focus on asymmetric multimodality. In the Monte Carlo experiment, we estimated 14 models using datasets created with a data generation process in WTP space. To ensure the stability of the parameter estimates, all models were estimated for 1,000 synthetic datasets and key conclusions were derived based on mean and modal values. The first objective of this study was to investigate the performance of LML models, with various numbers of parameters, on different workable sample sizes and with different panel length.

Our findings suggest that LML models require large sample sizes to outperform tradi-



Taste weekly, LML-Spline $\kappa = 22$

Taste never, LML-Spline $\kappa = 22$

Taste weekly, LML-Spline $\kappa = 44$

Odor never, LML-Spline $\kappa = 44$

Mild turbidity, LML-Spline $\kappa = 44$

Extreme turbidity, LML-Spline $\kappa = 44$

Figure 2: Distributions of WTP value estimates from various LML models.

tional MXL-N models. At sample sizes less than $N = 490$, LML models performed worse than the traditional specifications based on the assumption of normally distributed coefficients, unless the panel length $T = 8$. The second objective was to identify the optimal number of parameters to be adopted in LML model specification. Our hypothesis, based on previous findings of studies on flexible choice models (Fosgerau and Hess, 2007), was that increasing the number of parameters yields better approximations of the true distributions of the parameters.

Our empirical findings support this hypothesis with a practically important qualifier, in that LML specifications with large number of parameters outperformed those with small ones *only at large enough sample sizes*. At smaller sample size there seem to be a tradeoff between bias and variance in favour of MXL-N, especially under the correct utility specification (in this case in WTP space). The results from LML models estimated from datasets with low number of observations were mixed, and in many cases model specifications with small number of parameters outperformed those with larger number of parameters.

In our tap water preference study LML model suggests a pattern of multimodality that cannot be captured by the MNL-N or other unimodal parametric distributions. It is also inadequate to address such pattern with latent class models, as they do not produce a good fit and impose perfect correlation of random coefficients within classes, a restriction that the LML does not impose and for which we find no empirical evidence in our data. Regulators intending to achieve economically and politically efficient outcomes should be aware of the multimodal nature of preference for tap water in the tannery district of the Province of Vicenza. The tariff thresholds necessary to trigger majority voting in support of infrastructure investments that deliver only monthly chlorine smell in water and mild turbidity might be lower than those suggested by model estimates obtained with MXL-N models.

Overall, the results of our study do not support the blind use of very flexible mixing distributions, as at times LML models with a large number of parameters performed worse compared to both LML specifications with low number of parameters and MXL models. Thus, as a general guideline, we suggest to adopt LML model specifications with large number of parameters only when a sufficiently large number of observations is available, i.e. large sample sizes and adequately long panels.

While this study provides some insights about LML performance, additional simulation experiments are needed to evaluate the robustness of our conclusions. Additional experiments can include a variety of data settings such as variation in the number of alternatives, number of choice situations in the panel data, number of explanatory variables in the utility equation, and correlation among parameters. Importantly, on the practical side, analysts can no longer be excused to adopt parametric specifications without providing robust theoretical justifications corroborated by empirical evidence. This because the LML approach is sufficiently practical and general purpose software has been made available for all to use (Train, 2016) and it has been recently extended to allow for some fixed parameters in the specification (Bansal et al., 2016).

Tables

Table 1: MSE for means of random coefficients in DGP 1 (bimodal, $T = 4$)

Model	k	$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
		ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2
MXL-N Pref.	6	0.276	0.436	0.234	0.383	0.142	0.299	0.108	0.200	0.055	0.135
MXL-N WTP	6	0.165	0.301	0.095	0.281	0.038	0.208	0.013	0.128	0.009	0.104
LML-Poly	12	0.224	0.429	0.132	0.311	0.059	0.294	0.042	0.245	0.014	0.137
	24	0.268	0.404	0.104	0.359	0.064	0.225	0.022	0.125	0.009	0.077
	36	0.361	0.554	0.215	0.379	0.036	0.301	0.019	0.095	0.004	0.059
	48	0.276	0.395	0.212	0.349	0.037	0.229	0.012	0.097	0.005	0.054
LML-Step	12	0.245	0.407	0.237	0.312	0.039	0.231	0.054	0.248	0.023	0.160
	24	0.209	0.405	0.149	0.384	0.072	0.201	0.061	0.132	0.014	0.104
	36	0.212	0.326	0.174	0.304	0.094	0.225	0.026	0.101	0.012	0.056
	48	0.261	0.365	0.141	0.315	0.115	0.271	0.013	0.093	0.007	0.052
LML-Spline	12	0.288	0.485	0.243	0.322	0.069	0.235	0.089	0.219	0.021	0.084
	24	0.197	0.423	0.139	0.332	0.089	0.191	0.008	0.148	0.014	0.053
	36	0.263	0.456	0.201	0.463	0.126	0.231	0.022	0.128	0.006	0.049
	48	0.309	0.445	0.282	0.312	0.044	0.188	0.008	0.105	0.006	0.046

Table 2: MSE for standard deviations of random coefficients in DGP 1 (bimodal, $T = 4$)

Model	k	$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
		σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2
MXL-N Pref.	6	0.463	0.693	0.420	0.629	0.360	0.516	0.322	0.401	0.283	0.359
MXL-N WTP	6	0.370	0.542	0.319	0.516	0.293	0.475	0.286	0.376	0.246	0.346
LML-Poly	12	0.467	0.639	0.481	0.672	0.361	0.591	0.342	0.487	0.284	0.384
	24	0.484	0.628	0.341	0.551	0.351	0.487	0.265	0.329	0.245	0.297
	36	0.449	0.765	0.419	0.714	0.305	0.614	0.261	0.317	0.220	0.278
	48	0.411	0.645	0.329	0.561	0.267	0.495	0.246	0.293	0.214	0.267
LML-Step	12	0.392	0.597	0.421	0.554	0.341	0.579	0.363	0.469	0.252	0.373
	24	0.401	0.644	0.492	0.638	0.362	0.516	0.295	0.360	0.223	0.346
	36	0.424	0.553	0.425	0.505	0.330	0.440	0.255	0.333	0.222	0.289
	48	0.408	0.556	0.367	0.551	0.299	0.489	0.245	0.313	0.207	0.275
LML-Spline	12	0.421	0.591	0.775	0.649	0.312	0.421	0.305	0.449	0.282	0.378
	24	0.424	0.651	0.324	0.560	0.269	0.413	0.279	0.352	0.241	0.263
	36	0.467	0.718	0.449	0.624	0.334	0.582	0.262	0.342	0.232	0.288
	48	0.488	0.681	0.489	0.665	0.426	0.412	0.253	0.345	0.218	0.248

Table 3: MSE for means of random coefficients in DGP 1 (bimodal, $T = 8$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2
MXL-N Pref.	6	0.185	0.277	0.168	0.252	0.144	0.207	0.129	0.161	0.113	0.144
MXL-N WTP	6	0.148	0.217	0.128	0.207	0.117	0.190	0.115	0.151	0.099	0.139
LML-Poly	12	0.187	0.256	0.193	0.269	0.145	0.237	0.137	0.195	0.114	0.154
	24	0.194	0.251	0.137	0.221	0.141	0.195	0.106	0.132	0.098	0.119
	36	0.180	0.306	0.168	0.286	0.122	0.246	0.105	0.127	0.088	0.111
	48	0.165	0.258	0.132	0.225	0.107	0.198	0.099	0.117	0.086	0.107
LML-Step	12	0.157	0.239	0.169	0.222	0.137	0.232	0.145	0.188	0.101	0.149
	24	0.161	0.258	0.197	0.255	0.145	0.207	0.118	0.144	0.085	0.139
	36	0.170	0.221	0.170	0.202	0.132	0.176	0.102	0.133	0.089	0.116
	48	0.163	0.223	0.147	0.221	0.120	0.196	0.098	0.125	0.083	0.110
LML-Spline	12	0.169	0.237	0.210	0.260	0.125	0.169	0.122	0.180	0.113	0.151
	24	0.170	0.261	0.130	0.224	0.108	0.165	0.112	0.141	0.097	0.105
	36	0.187	0.287	0.180	0.250	0.134	0.233	0.105	0.137	0.093	0.115
	48	0.195	0.273	0.196	0.266	0.171	0.165	0.101	0.138	0.087	0.099

Table 4: MSE for standard deviations of random coefficients in DGP 1 (bimodal, $T = 8$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2
MXL-N Pref.	6	0.204	0.305	0.185	0.277	0.158	0.228	0.142	0.177	0.124	0.158
MXL-N WTP	6	0.163	0.239	0.141	0.228	0.129	0.209	0.127	0.166	0.109	0.153
LML-Poly	12	0.206	0.282	0.212	0.296	0.160	0.261	0.151	0.215	0.125	0.169
	24	0.213	0.276	0.151	0.243	0.155	0.215	0.117	0.145	0.108	0.131
	36	0.198	0.337	0.185	0.315	0.134	0.271	0.116	0.140	0.097	0.122
	48	0.182	0.284	0.145	0.248	0.118	0.218	0.109	0.129	0.095	0.118
LML-Step	12	0.173	0.263	0.186	0.244	0.151	0.255	0.160	0.207	0.111	0.164
	24	0.177	0.284	0.217	0.281	0.160	0.228	0.130	0.158	0.094	0.153
	36	0.187	0.243	0.187	0.222	0.145	0.194	0.112	0.146	0.098	0.128
	48	0.179	0.245	0.162	0.243	0.132	0.216	0.108	0.138	0.091	0.121
LML-Spline	12	0.186	0.261	0.231	0.286	0.138	0.186	0.134	0.198	0.124	0.166
	24	0.187	0.287	0.143	0.246	0.119	0.182	0.123	0.155	0.107	0.116
	36	0.206	0.316	0.198	0.275	0.147	0.256	0.116	0.151	0.102	0.127
	48	0.215	0.300	0.216	0.293	0.188	0.182	0.111	0.152	0.096	0.109

Table 5: MSE for means of random coefficients in DGP 2 (trimodal, $T = 4$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2
MXL-N Pref.	6	0.248	0.392	0.211	0.345	0.128	0.269	0.097	0.180	0.050	0.122
MXL-N WTP	6	0.149	0.271	0.086	0.253	0.074	0.187	0.068	0.115	0.008	0.094
LML-Poly	12	0.202	0.386	0.119	0.280	0.053	0.265	0.038	0.221	0.013	0.123
	24	0.241	0.364	0.094	0.323	0.058	0.203	0.020	0.113	0.008	0.069
	36	0.325	0.499	0.194	0.341	0.032	0.271	0.017	0.086	0.004	0.053
	48	0.248	0.356	0.191	0.314	0.033	0.206	0.011	0.087	0.005	0.049
LML-Step	12	0.221	0.366	0.213	0.281	0.035	0.208	0.049	0.223	0.021	0.144
	24	0.188	0.365	0.134	0.346	0.065	0.181	0.035	0.119	0.013	0.094
	36	0.191	0.293	0.157	0.274	0.085	0.203	0.023	0.091	0.011	0.050
	48	0.235	0.329	0.127	0.284	0.104	0.244	0.012	0.084	0.006	0.047
LML-Spline	12	0.259	0.437	0.219	0.290	0.062	0.212	0.080	0.197	0.019	0.076
	24	0.177	0.381	0.125	0.299	0.080	0.172	0.057	0.133	0.013	0.048
	36	0.237	0.410	0.181	0.417	0.113	0.208	0.020	0.115	0.005	0.044
	48	0.278	0.401	0.254	0.281	0.040	0.169	0.007	0.095	0.005	0.041

Table 6: MSE for standard deviations of random coefficients in DGP 2 (trimodal, $T = 4$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2
MXL-N Pref.	6	0.426	0.633	0.392	0.581	0.341	0.477	0.352	0.368	0.304	0.342
MXL-N WTP	6	0.371	0.500	0.301	0.479	0.281	0.443	0.344	0.398	0.294	0.328
LML-Poly	12	0.432	0.592	0.447	0.614	0.340	0.549	0.312	0.469	0.283	0.356
	24	0.452	0.577	0.319	0.510	0.329	0.447	0.271	0.367	0.262	0.323
	36	0.413	0.701	0.388	0.654	0.288	0.566	0.216	0.369	0.262	0.257
	48	0.383	0.596	0.307	0.522	0.255	0.462	0.230	0.304	0.244	0.250
LML-Step	12	0.368	0.551	0.389	0.511	0.318	0.538	0.365	0.424	0.256	0.306
	24	0.373	0.592	0.457	0.584	0.337	0.482	0.271	0.357	0.233	0.290
	36	0.394	0.507	0.392	0.470	0.315	0.410	0.272	0.374	0.223	0.268
	48	0.382	0.515	0.342	0.506	0.282	0.458	0.272	0.332	0.192	0.262
LML-Spline	12	0.395	0.542	0.707	0.600	0.293	0.391	0.339	0.452	0.289	0.281
	24	0.394	0.597	0.306	0.517	0.252	0.383	0.290	0.399	0.280	0.271
	36	0.432	0.659	0.418	0.576	0.312	0.540	0.315	0.383	0.285	0.262
	48	0.453	0.626	0.456	0.610	0.401	0.385	0.280	0.326	0.247	0.257

Table 7: MSE for means of random coefficients in DGP 2 (trimodal, $T = 8$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2	ω^1	ω^2
MXL-N Pref.	6	0.167	0.249	0.151	0.227	0.130	0.186	0.116	0.145	0.102	0.130
MXL-N WTP	6	0.133	0.195	0.115	0.186	0.105	0.171	0.104	0.136	0.089	0.125
LML-Poly	12	0.168	0.230	0.174	0.242	0.131	0.213	0.123	0.176	0.103	0.139
	24	0.175	0.226	0.123	0.199	0.127	0.176	0.095	0.119	0.088	0.107
	36	0.162	0.275	0.151	0.257	0.110	0.221	0.095	0.114	0.079	0.100
	48	0.149	0.232	0.119	0.203	0.096	0.178	0.089	0.105	0.077	0.096
LML-Step	12	0.141	0.215	0.152	0.200	0.123	0.209	0.131	0.169	0.091	0.134
	24	0.145	0.232	0.177	0.230	0.131	0.186	0.106	0.130	0.077	0.125
	36	0.153	0.199	0.153	0.182	0.119	0.158	0.092	0.120	0.080	0.104
	48	0.147	0.201	0.132	0.199	0.108	0.176	0.088	0.113	0.075	0.099
LML-Spline	12	0.152	0.213	0.189	0.234	0.113	0.152	0.110	0.162	0.102	0.136
	24	0.153	0.235	0.117	0.202	0.097	0.149	0.101	0.127	0.087	0.095
	36	0.168	0.258	0.162	0.225	0.121	0.210	0.095	0.123	0.084	0.104
	48	0.176	0.246	0.176	0.239	0.154	0.149	0.091	0.124	0.078	0.089

Table 8: MSE for standard deviations of random coefficients in DGP 2 (trimodal, $T = 4$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2	σ_1	σ_2
MXL-N Pref.	6	0.184	0.275	0.167	0.249	0.142	0.205	0.128	0.159	0.112	0.142
MXL-N WTP	6	0.147	0.215	0.127	0.205	0.116	0.188	0.114	0.149	0.098	0.138
LML-Poly	12	0.185	0.254	0.191	0.266	0.144	0.235	0.136	0.194	0.113	0.152
	24	0.192	0.248	0.136	0.219	0.140	0.194	0.105	0.131	0.097	0.118
	36	0.178	0.303	0.167	0.284	0.121	0.244	0.104	0.126	0.087	0.110
	48	0.164	0.256	0.131	0.223	0.106	0.196	0.098	0.116	0.086	0.106
LML-Step	12	0.156	0.237	0.167	0.220	0.136	0.230	0.144	0.186	0.100	0.148
	24	0.159	0.256	0.195	0.253	0.144	0.205	0.117	0.142	0.085	0.138
	36	0.168	0.219	0.168	0.200	0.131	0.175	0.101	0.131	0.088	0.115
	48	0.161	0.221	0.146	0.219	0.119	0.194	0.097	0.124	0.082	0.109
LML-Spline	12	0.167	0.235	0.208	0.257	0.124	0.167	0.121	0.178	0.112	0.149
	24	0.168	0.258	0.129	0.221	0.107	0.164	0.111	0.140	0.096	0.104
	36	0.185	0.284	0.178	0.248	0.132	0.230	0.104	0.136	0.092	0.114
	48	0.194	0.270	0.194	0.264	0.169	0.164	0.100	0.137	0.086	0.098

Table 9: Means and st. dev. of modal estimates of ω_n^1 in DGP 1 (bimodal, $T = 4$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**
Real	-	0.55	1.26	0.55	1.26	0.55	1.26	0.55	1.26	0.55	1.26
MXL−N Pref.	6	0.89	–	0.84	–	0.75	–	0.86	–	0.80	–
		(0.11)	–	(0.14)	–	(0.10)	–	(0.15)	–	(0.16)	–
MXL−N WTP	6	0.99	–	1.02	–	1.11	–	0.94	–	0.84	–
		(0.14)	–	(0.18)	–	(0.19)	–	(0.21)	–	(0.17)	–
LML-Poly	12	0.68	2.06	0.71	1.99	0.74	1.77	0.65	1.56	0.42	1.36
		(0.07)	(0.24)	(0.06)	(0.21)	(0.08)	(0.16)	(0.05)	(0.15)	(0.03)	(0.12)
	24	0.71	2.11	0.68	2.08	0.74	1.63	0.63	1.44	0.60	1.34
		(0.07)	(0.31)	(0.06)	(0.22)	(0.06)	(0.15)	(0.06)	(0.20)	(0.03)	(0.10)
	36	0.69	2.07	0.61	1.99	0.61	1.69	0.47	1.49	0.55	1.25
		(0.05)	(0.20)	(0.06)	(0.19)	(0.04)	(0.18)	(0.04)	(0.17)	(0.03)	(0.11)
	48	0.74	2.08	0.66	2.02	0.62	1.66	0.49	1.45	0.52	1.22
		(0.06)	(0.22)	(0.05)	(0.21)	(0.07)	(0.14)	(0.04)	(0.14)	(0.02)	(0.09)
LML-Step	12	0.77	2.05	0.74	2.03	0.75	1.80	0.67	1.45	0.44	1.35
		(0.07)	(0.22)	(0.06)	(0.21)	(0.09)	(0.19)	(0.05)	(0.18)	(0.05)	(0.14)
	24	0.79	2.09	0.72	2.05	0.71	1.67	0.62	1.40	0.49	1.39
		(0.07)	(0.26)	(0.05)	(0.27)	(0.08)	(0.16)	(0.03)	(0.16)	(0.03)	(0.11)
	36	0.76	2.01	0.61	1.92	0.61	1.50	0.49	1.37	0.53	1.22
		(0.07)	(0.25)	(0.09)	(0.21)	(0.07)	(0.18)	(0.04)	(0.18)	(0.04)	(0.09)
	48	0.71	2.06	0.64	1.93	0.60	1.53	0.48	1.32	0.54	1.29
		(0.06)	(0.19)	(0.06)	(0.22)	(0.06)	(0.16)	(0.06)	(0.14)	(0.03)	(0.08)
LML-Spline	12	0.78	2.13	0.63	2.10	0.66	1.52	0.68	1.44	0.53	1.32
		(0.09)	(0.26)	(0.07)	(0.20)	(0.08)	(0.21)	(0.06)	(0.18)	(0.05)	(0.11)
	24	0.76	2.19	0.65	2.06	0.72	1.56	0.67	1.39	0.54	1.31
		(0.08)	(0.21)	(0.08)	(0.23)	(0.08)	(0.20)	(0.06)	(0.19)	(0.05)	(0.12)
	36	0.76	2.13	0.67	2.08	0.65	1.49	0.48	1.34	0.58	1.24
		(0.07)	(0.32)	(0.08)	(0.24)	(0.08)	(0.19)	(0.06)	(0.20)	(0.05)	(0.08)
	48	0.75	2.18	0.71	2.08	0.62	1.44	0.46	1.31	0.53	1.21
		(0.07)	(0.24)	(0.05)	(0.21)	(0.08)	(0.19)	(0.06)	(0.17)	(0.05)	(0.08)

Table 10: Means and st. dev. of modal estimates of ω_n^1 in DGP 1 (bimodal, $T = 8$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**
Real	-	0.55	1.26	0.55	1.26	0.55	1.26	0.55	1.26	0.55	1.26
MXL−N Pref.	6	0.88	–	0.80	–	0.77	–	0.74	–	0.71	–
		(0.14)	–	(0.12)	–	(0.12)	–	(0.18)	–	(0.15)	–
MXL−N WTP	6	0.95	–	0.93	–	0.91		0.94	–	0.88	–
		(0.20)	–	(0.15)	–	(0.12)	–	(0.18)	–	(0.14)	–
LML-Poly	12	0.71	2.14	0.74	2.07	0.67	1.65	0.67	1.48	0.64	1.43
		(0.06)	(0.22)	(0.06)	(0.15)	(0.07)	(0.18)	(0.05)	(0.14)	(0.04)	(0.11)
	24	0.74	2.19	0.71	2.16	0.66	1.51	0.65	1.47	0.62	1.41
		(0.06)	(0.21)	(0.07)	(0.19)	(0.05)	(0.18)	(0.06)	(0.17)	(0.06)	(0.12)
	36	0.72	2.15	0.73	2.07	0.73	1.57	0.69	1.42	0.56	1.32
		(0.07)	(0.19)	(0.05)	(0.17)	(0.06)	(0.17)	(0.04)	(0.15)	(0.05)	(0.09)
	48	0.77	2.16	0.69	2.10	0.64	1.54	0.61	1.39	0.54	1.29
		(0.05)	(0.18)	(0.06)	(0.16)	(0.05)	(0.15)	(0.04)	(0.13)	(0.04)	(0.08)
LML-Step	12	0.70	2.13	0.77	2.11	0.77	1.78	0.69	1.42	0.66	1.42
		(0.11)	(0.22)	(0.14)	(0.22)	(0.12)	(0.19)	(0.09)	(0.12)	(0.05)	(0.12)
	24	0.72	2.17	0.75	2.13	0.73	1.75	0.64	1.37	0.51	1.37
		(0.10)	(0.24)	(0.09)	(0.22)	(0.08)	(0.15)	(0.09)	(0.11)	(0.06)	(0.10)
	36	0.69	2.08	0.73	2.00	0.63	1.78	0.61	1.39	0.55	1.29
		(0.12)	(0.22)	(0.09)	(0.21)	(0.08)	(0.14)	(0.08)	(0.12)	(0.05)	(0.09)
	48	0.73	2.14	0.67	2.01	0.64	1.71	0.60	1.36	0.54	1.26
		(0.09)	(0.19)	(0.06)	(0.22)	(0.07)	(0.15)	(0.06)	(0.11)	(0.04)	(0.08)
LML-Spline	12	0.71	2.22	0.66	2.18	0.68	1.80	0.60	1.44	0.63	1.44
		(0.09)	(0.22)	(0.08)	(0.20)	(0.08)	(0.21)	(0.05)	(0.20)	(0.05)	(0.15)
	24	0.69	2.28	0.77	2.14	0.74	1.84	0.69	1.48	0.59	1.48
		(0.08)	(0.24)	(0.09)	(0.19)	(0.08)	(0.20)	(0.06)	(0.18)	(0.06)	(0.11)
	36	0.69	2.22	0.72	2.16	0.67	1.70	0.60	1.33	0.57	1.31
		(0.06)	(0.29)	(0.10)	(0.24)	(0.09)	(0.17)	(0.06)	(0.21)	(0.05)	(0.08)
	48	0.62	2.27	0.74	2.16	0.64	1.75	0.58	1.38	0.55	1.28
		(0.05)	(0.16)	(0.07)	(0.21)	(0.06)	(0.15)	(0.04)	(0.10)	(0.03)	(0.07)

Table 11: Means and st. dev. of modal estimates of ω_n^2 in DGP 1 (bimodal, $T = 4$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**
Real	-	-1.78	1.69	-1.78	1.69	-1.78	1.69	-1.78	1.69	-1.78	1.69
MXL−N Pref.	6	0.78	–	0.76	–	0.65	–	0.56	–	0.46	–
		(0.16)	–	(0.15)	–	(0.11)	–	(0.12)	–	(0.09)	–
MXL−N WTP	6	0.55	–	0.43	–	0.38		0.29	–	0.18	–
		(0.18)	–	(0.11)	–	(0.10)	–	(0.08)	–	(0.04)	–
LML-Poly	12	-1.14	1.22	-1.19	1.18	-1.18	1.41	-1.22	1.43	-1.23	1.50
		(0.16)	(0.25)	(0.14)	(0.22)	(0.14)	(0.16)	(0.15)	(0.19)	(0.13)	(0.15)
	24	-1.19	1.35	-1.20	1.26	-1.34	1.31	-1.37	1.37	-1.49	1.40
		(0.18)	(0.24)	(0.12)	(0.21)	(0.12)	(0.17)	(0.12)	(0.16)	(0.12)	(0.12)
	36	-1.29	1.42	-1.28	1.45	-1.40	1.44	-1.41	1.54	-1.51	1.55
		(0.11)	(0.20)	(0.12)	(0.18)	(0.11)	(0.12)	(0.17)	(0.16)	(0.15)	(0.16)
	48	-1.38	1.43	-1.45	1.51	-1.53	1.52	-1.49	1.54	-1.64	1.62
		(0.11)	(0.14)	(0.11)	(0.14)	(0.09)	(0.11)	(0.08)	(0.11)	(0.08)	(0.07)
LML-Step	12	-1.20	1.49	-1.16	1.46	-1.26	1.12	-1.33	1.15	-1.43	1.20
		(0.14)	(0.25)	(0.15)	(0.20)	(0.14)	(0.19)	(0.15)	(0.15)	(0.14)	(0.16)
	24	-1.37	1.49	-1.43	1.57	-1.37	1.34	-1.46	1.44	-1.50	1.42
		(0.16)	(0.21)	(0.13)	(0.13)	(0.17)	(0.18)	(0.14)	(0.12)	(0.13)	(0.14)
	36	-1.40	1.36	-1.46	1.40	-1.50	1.45	-1.59	1.46	-1.67	1.51
		(0.14)	(0.17)	(0.12)	(0.12)	(0.16)	(0.13)	(0.17)	(0.09)	(0.12)	(0.11)
	48	-1.26	1.11	-1.32	1.14	-1.35	1.49	-1.42	1.61	-1.66	1.73
		(0.15)	(0.15)	(0.16)	(0.15)	(0.13)	(0.12)	(0.09)	(0.09)	(0.08)	(0.06)
LML-Spline	12	-1.28	1.19	-1.23	1.12	-1.39	1.21	-1.42	1.33	-1.46	1.41
		(0.15)	(0.22)	(0.17)	(0.18)	(0.15)	(0.16)	(0.15)	(0.19)	(0.12)	(0.14)
	24	-1.28	1.19	-1.23	1.17	-1.51	1.22	-1.60	1.38	-1.54	1.40
		(0.15)	(0.18)	(0.15)	(0.21)	(0.12)	(0.16)	(0.14)	(0.16)	(0.13)	(0.16)
	36	-1.36	1.03	-1.42	1.16	-1.33	1.55	-1.40	1.59	-1.54	1.66
		(0.16)	(0.20)	(0.18)	(0.19)	(0.1)	(0.16)	(0.12)	(0.11)	(0.10)	(0.12)
	48	-1.32	1.12	-1.31	1.17	-1.43	1.59	-1.55	1.53	-1.66	1.70
		(0.16)	(0.16)	(0.14)	(0.15)	(0.09)	(0.10)	(0.08)	(0.12)	(0.07)	(0.09)

Table 12: Means and st. dev. of modal estimates of ω_n^2 in DGP 1 (bimodal, $T = 8$)

Model		$N = 70$		$N = 210$		$N = 490$		$N = 980$		$N = 1,960$	
	k	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**	**Max1**	**Max2**
Real	–	−1.78	1.69	−1.78	1.69	−1.78	1.69	−1.78	1.69	−1.78	1.69
MXL−N Pref.	6	0.22	–	0.34	–	0.27	–	0.49	–	0.45	–
		(0.04)	–	(0.05)	–	(0.02)	–	(0.08)	–	(0.05)	–
MXL−N WTP	6	0.12	–	0.12	–	0.17		0.27	–	0.08	–
		(0.03)	–	(0.05)	–	(0.04)	–	(0.06)	–	(0.03)	–
LML-Poly	12	−1.19	1.27	−1.24	1.23	−1.23	1.47	−1.27	1.49	−1.28	1.56
		(0.19)	(0.25)	(0.16)	(0.20)	(0.15)	(0.18)	(0.16)	(0.15)	(0.12)	(0.14)
	24	−1.24	1.40	−1.25	1.31	−1.39	1.36	−1.42	1.42	−1.55	1.46
		(0.16)	(0.22)	(0.11)	(0.18)	(0.14)	(0.15)	(0.14)	(0.14)	(0.11)	(0.12)
	36	−1.34	1.48	−1.33	1.51	−1.46	1.50	−1.47	1.60	−1.57	1.61
		(0.11)	(0.21)	(0.14)	(0.15)	(0.12)	(0.13)	(0.15)	(0.13)	(0.10)	(0.13)
	48	−1.44	1.49	−1.51	1.57	−1.59	1.58	−1.55	1.60	−1.71	1.68
		(0.13)	(0.12)	(0.10)	(0.13)	(0.09)	(0.10)	(0.09)	(0.10)	(0.06)	(0.08)
LML-Step	12	−1.25	1.55	−1.21	1.52	−1.31	1.16	−1.38	1.20	−1.49	1.25
		(0.20)	(0.25)	(0.13)	(0.12)	(0.14)	(0.19)	(0.14)	(0.12)	(0.14)	(0.16)
	24	−1.42	1.55	−1.49	1.63	−1.42	1.39	−1.52	1.50	−1.56	1.48
		(0.11)	(0.21)	(0.13)	(0.13)	(0.17)	(0.18)	(0.14)	(0.12)	(0.12)	(0.16)
	36	−1.46	1.41	−1.52	1.46	−1.56	1.51	−1.65	1.52	−1.74	1.57
		(0.13)	(0.17)	(0.12)	(0.12)	(0.18)	(0.13)	(0.17)	(0.09)	(0.12)	(0.14)
	48	−1.31	1.15	−1.37	1.19	−1.40	1.55	−1.48	1.67	−1.73	1.80
		(0.12)	(0.13)	(0.15)	(0.15)	(0.11)	(0.10)	(0.09)	(0.11)	(0.06)	(0.10)
LML-Spline	12	−1.33	1.24	−1.28	1.16	−1.45	1.26	−1.48	1.38	−1.52	1.47
		(0.18)	(0.22)	(0.19)	(0.18)	(0.15)	(0.16)	(0.15)	(0.19)	(0.12)	(0.14)
	24	−1.37	1.28	−1.24	1.22	−1.57	1.27	−1.66	1.44	−1.60	1.46
		(0.15)	(0.18)	(0.15)	(0.18)	(0.12)	(0.16)	(0.14)	(0.16)	(0.13)	(0.16)
	36	−1.41	1.07	−1.48	1.21	−1.38	1.61	−1.46	1.65	−1.60	1.73
		(0.16)	(0.17)	(0.11)	(0.19)	(0.1)	(0.19)	(0.12)	(0.11)	(0.10)	(0.12)
	48	−1.37	1.16	−1.36	1.22	−1.49	1.65	−1.61	1.59	−1.73	1.77
		(0.14)	(0.11)	(0.12)	(0.11)	(0.09)	(0.09)	(0.07)	(0.12)	(0.07)	(0.06)

Table 13: Means and st. dev. of modal estimates of ω_n^1 in DGP 2 (trimodal, $T = 4$)

Model		$N = 70$			$N = 210$			$N = 490$			$N = 980$			$N = 1,960$		
	k	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**
Real		−1.12	1.77	3.61	−1.12	1.77	3.61	−1.12	1.77	3.61	−1.12	1.77	3.61	−1.12	1.77	3.61
MXL−N Pref.	6	1.13	–	–	1.34	–	–	1.27	–	–	1.19	–	–	1.25	–	–
		(0.26)			(0.29)			(0.28)			(0.23)			(0.22)		
MXL−N WTP	6	1.15	–	–	1.13	–	–	1.18	–	–	1.27	–	–	1.38	–	–
		(0.26)			(0.19)			(0.27)			(0.28)			(0.20)		
LML-Poly	12	−1.89	3.10	–	−1.80	3.16	–	−1.76	3.23	–	−1.72	3.29	–	−1.38	3.36	–
		(0.21)	(0.35)		(0.22)	(0.18)		(0.19)	(0.29)		(0.15)	(0.26)		(0.13)	(0.25)	
	24	−1.91	3.19	–	−1.81	3.25	–	−1.78	3.32	–	−1.74	2.22	4.04	−1.39	2.00	3.86
		(0.22)	(0.37)		(0.24)	(0.19)		(0.22)	(0.25)		(0.14)	(0.28)	(0.39)	(0.15)	(0.25)	(0.32)
	36	−1.86	3.32	–	−1.77	2.24	4.37	−1.73	2.20	4.28	−1.70	2.15	4.20	−1.36	1.94	3.72
		(0.19)	(0.32)		(0.21)	(0.25)	(0.41)	(0.24)	(0.24)	(0.38)	(0.13)	(0.21)	(0.35)	(0.16)	(0.22)	(0.34)
	48	−1.77	2.33	4.39	−1.68	2.21	4.17	−1.79	2.17	4.09	−1.61	2.13	4.01	−1.29	1.91	3.68
		(0.18)	(0.28)	(0.46)	(0.19)	(0.26)	(0.38)	(0.22)	(0.21)	(0.36)	(0.10)	(0.20)	(0.33)	(0.12)	(0.18)	(0.29)
LML-Step	12	−1.98	3.26	–	−1.88	3.32	–	−1.84	3.39	–	−1.81	3.45	–	−1.45	3.52	–
		(0.24)	(0.31)		(0.21)	(0.22)		(0.22)	(0.19)		(0.12)	(0.29)		(0.12)	(0.17)	
	24	−1.90	3.35	–	−1.80	3.42	–	−1.76	3.48	–	−1.73	2.19	4.15	−1.38	1.97	3.82
		(0.23)	(0.27)		(0.20)	(0.22)		(0.23)	(0.18)		(0.15)	(0.31)	(0.46)	(0.13)	(0.16)	(0.33)
	36	−1.83	3.49	–	−1.74	2.26	4.20	−1.70	2.22	4.37	−1.67	2.17	4.03	−1.34	1.95	3.71
		(0.25)	(0.26)		(0.24)	(0.25)	(0.42)	(0.27)	(0.16)	(0.39)	(0.14)	(0.26)	(0.44)	(0.15)	(0.14)	(0.36)
	48	−1.76	2.42	4.46	−1.67	2.30	4.24	−1.77	2.25	4.24	−1.61	2.21	4.07	−1.28	1.89	3.68
		(0.21)	(0.22)	(0.51)	(0.21)	(0.23)	(0.40)	(0.19)	(0.20)	(0.37)	(0.14)	(0.20)	(0.39)	(0.09)	(0.15)	(0.34)
LML-Spline	12	−2.01	3.18	–	−1.91	3.24	–	−1.87	3.31	–	−1.83	3.37	–	−1.47	3.44	–
		(0.14)	(0.27)		(0.21)	(0.19)		(0.22)	(0.21)		(0.19)	(0.15)		(0.16)	(0.25)	
	24	−1.89	3.27	–	−1.90	3.34	–	−1.76	3.40	–	−1.75	2.15	4.11	−1.38	2.01	3.79
		(0.15)	(0.29)		(0.19)	(0.24)		(0.22)	(0.23)		(0.18)	(0.25)	(0.49)	(0.14)	(0.29)	(0.40)
	36	−1.89	3.40	–	−1.80	3.19	–	−1.79	3.15	–	−1.72	2.11	4.09	−1.35	1.93	3.76
		(0.12)	(0.32)		(0.24)	(0.15)		(0.22)	(0.24)		(0.17)	(0.26)	(0.42)	(0.16)	(0.28)	(0.42)
	48	−1.83	2.39	4.40	−1.74	2.27	4.18	−1.70	2.23	4.10	−1.71	2.18	4.01	−1.34	1.92	3.69
		(0.16)	(0.24)	(0.44)	(0.20)	(0.13)	(0.41)	(0.22)	(0.17)	(0.22)	(0.11)	(0.23)	(0.38)	(0.11)	(0.24)	(0.36)

Table 14: Means and st. dev. of modal estimates of ω_n^1 in DGP 2 (trimodal, $T = 8$)

Model		$N = 70$			$N = 210$			$N = 490$			$N = 980$			$N = 1{,}960$		
	k	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**
Real		−1.12	1.77	3.61	−1.12	1.77	3.61	−1.12	1.77	3.61	−1.12	1.77	3.61	−1.12	1.77	3.61
MXL−N Pref.	6	1.24	–	–	1.28	–	–	1.35	–	–	1.38	–	–	1.43	–	–
		(0.22)			(0.23)			(0.25)			(0.23)			(0.28)		
MXL−N WTP	6	1.26	–	–	1.33	–	–	1.38	–	–	1.57	–	–	1.49	–	–
		(0.27)			(0.30)			(0.25)			(0.26)			(0.22)		
LML-Poly	12	−1.80	2.95	–	−1.71	3.00	–	−1.67	3.06	–	−1.64	3.13	–	−1.31	3.19	–
		(0.23)	(0.33)		(0.21)	(0.19)		(0.20)	(0.31)		(0.13)	(0.26)		(0.14)	(0.22)	
	24	−1.81	3.03	–	−1.72	3.09	–	−1.69	3.15	4.18	−1.66	2.11	3.84	−1.32	1.90	3.67
		(0.25)	(0.35)		(0.24)	(0.21)		(0.23)	(0.23)	(0.41)	(0.12)	(0.28)	(0.37)	(0.13)	(0.26)	(0.35)
	36	−1.77	3.15	–	−1.68	2.13	4.15	−1.65	2.09	4.07	−1.61	2.05	3.99	−1.29	1.84	3.53
		(0.20)	(0.30)		(0.21)	(0.24)	(0.42)	(0.22)	(0.22)	(0.36)	(0.12)	(0.21)	(0.36)	(0.14)	(0.24)	(0.33)
	48	−1.68	2.21	4.17	−1.60	2.10	3.96	−1.57	2.06	3.88	−1.53	2.02	3.81	−1.23	1.82	3.50
		(0.16)	(0.22)	(0.44)	(0.20)	(0.24)	(0.37)	(0.21)	(0.19)	(0.38)	(0.12)	(0.20)	(0.32)	(0.11)	(0.15)	(0.28)
LML-Step	12	−1.88	3.09	–	−1.79	3.15	–	−1.75	3.22	–	−1.72	3.28	–	−1.37	3.35	–
		(0.25)	(0.30)		(0.20)	(0.23)		(0.25)	(0.19)		(0.11)	(0.23)		(0.13)	(0.17)	
	24	−1.80	3.18	–	−1.71	3.25	–	−1.68	3.31	4.21	−1.64	2.08	3.94	−1.31	1.87	3.63
		(0.24)	(0.24)		(0.22)	(0.24)		(0.26)	(0.17)	(0.43)	(0.15)	(0.32)	(0.48)	(0.13)	(0.14)	(0.37)
	36	−1.74	3.31	–	−1.65	2.15	3.99	−1.62	2.10	4.16	−1.59	2.06	3.91	−1.27	1.86	3.58
		(0.22)	(0.25)		(0.25)	(0.22)	(0.45)	(0.23)	(0.18)	(0.37)	(0.14)	(0.22)	(0.43)	(0.14)	(0.14)	(0.38)
	48	−1.67	2.30	4.24	−1.68	2.18	4.03	−1.66	2.14	4.08	−1.53	2.02	3.87	−1.22	1.81	3.51
		(0.24)	(0.24)	(0.53)	(0.22)	(0.25)	(0.42)	(0.21)	(0.22)	(0.34)	(0.17)	(0.19)	(0.40)	(0.10)	(0.16)	(0.33)
LML-Spline	12	−1.91	3.02	–	−1.81	3.08	–	−1.78	3.14	–	−1.74	3.20	–	−1.39	3.27	–
		(0.16)	(0.27)		(0.21)	(0.19)		(0.22)	(0.21)		(0.19)	(0.15)		(0.16)	(0.25)	
	24	−1.80	3.11	–	−1.71	3.17	–	−1.67	3.23	4.32	−1.64	2.05	3.91	−1.31	1.84	3.60
		(0.15)	(0.29)		(0.19)	(0.24)		(0.22)	(0.23)	(0.46)	(0.18)	(0.25)	(0.49)	(0.14)	(0.29)	(0.40)
	36	−1.88	3.23	–	−1.73	2.08	4.12	−1.67	2.04	4.04	−1.66	2.00	3.88	−1.29	1.83	3.57
		(0.15)	(0.32)		(0.24)	(0.15)	(0.48)	(0.22)	(0.24)	(0.44)	(0.17)	(0.26)	(0.42)	(0.16)	(0.28)	(0.42)
	48	−1.74	2.27	4.18	−1.65	2.16	3.97	−1.62	2.11	3.96	−1.59	1.98	3.81	−1.27	1.79	3.51
		(0.15)	(0.24)	(0.44)	(0.20)	(0.13)	(0.41)	(0.22)	(0.17)	(0.22)	(0.11)	(0.23)	(0.38)	(0.11)	(0.24)	(0.36)

Table 15: Means and st. dev. of modal estimates of ω_n^2 in DGP 2 (trimodal, $T = 4$)

Model		$N = 70$			$N = 210$			$N = 490$			$N = 980$			$N = 1{,}960$		
	k	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**
Real		1.36	3.68	5.82	1.36	3.68	5.82	1.36	3.68	5.82	1.36	3.68	5.82	1.36	3.68	5.82
MXL−N Pref.	6	3.92	–	–	4.01	–	–	4.11	–	–	4.05	–	–	4.01	–	–
		(0.47)			(0.51)			(0.48)			(0.46)			(0.42)		
MXL−N WTP	6	4.10	–	–	4.04	–	–	3.99	–	–	4.08	–	–	4.12	–	–
		(0.58)			(0.49)			(0.47)			(0.39)			(0.36)		
LML-Poly	12	0.26	3.24	–	0.41	3.21	–	0.74	3.46	–	0.80	3.53	–	0.87	3.60	–
		(0.04)	(0.41)		(0.05)	(0.28)		(0.13)	(0.29)		(0.08)	(0.30)		(0.12)	(0.29)	
	24	0.25	3.21	–	0.24	3.60	–	0.29	3.68	–	0.31	3.28	5.25	0.34	3.38	5.34
		(0.04)	(0.36)		(0.06)	(0.31)		(0.09)	(0.32)		(0.05)	(0.29)	(0.48)	(0.03)	(0.28)	(0.45)
	36	0.29	3.29	–	0.50	3.08	5.14	0.67	3.31	5.21	0.72	3.38	5.31	0.79	3.45	5.43
		(0.03)	(0.37)		(0.11)	(0.30)	(0.55)	(0.16)	(0.26)	(0.51)	(0.11)	(0.28)	(0.49)	(0.10)	(0.28)	(0.42)
	48	0.68	3.01	5.15	0.79	3.32	5.21	0.87	3.34	5.27	0.94	3.41	5.34	1.02	3.50	5.55
		(0.13)	(0.35)	(0.53)	(0.12)	(0.29)	(0.48)	(0.19)	(0.28)	(0.46)	(0.08)	(0.25)	(0.44)	(0.08)	(0.25)	(0.39)
LML-Step	12	0.33	3.31	–	0.38	3.50	–	0.48	3.63	–	0.51	3.70	–	0.56	3.75	–
		(0.05)	(0.40)		(0.06)	(0.30)		(0.07)	(0.32)		(0.06)	(0.26)		(0.08)	(0.26)	
	24	0.15	3.32	–	0.17	3.28	–	0.17	3.24	–	0.19	3.25	5.16	0.21	3.27	5.29
		(0.02)	(0.38)		(0.03)	(0.31)		(0.01)	(0.29)		(0.02)	(0.25)	(0.46)	(0.05)	(0.30)	(0.44)
	36	0.22	3.38	–	0.43	3.12	5.18	0.69	3.19	5.20	0.75	3.30	5.22	0.81	3.37	5.25
		(0.02)	(0.35)		(0.08)	(0.28)	(0.43)	(0.10)	(0.28)	(0.47)	(0.12)	(0.27)	(0.42)	(0.10)	(0.28)	(0.45)
	48	0.78	3.04	5.22	0.82	3.22	5.25	0.85	3.29	5.23	0.92	3.36	5.31	1.00	3.45	5.49
		(0.10)	(0.32)	(0.49)	(0.11)	(0.29)	(0.45)	(0.12)	(0.25)	(0.41)	(0.15)	(0.24)	(0.40)	(0.13)	(0.22)	(0.38)
LML-Spline	12	0.40	3.36	–	0.61	3.47	–	0.91	3.69	–	0.98	3.77	–	1.07	3.86	–
		(0.06)	(0.34)		(0.25)	(0.20)		(0.24)	(0.35)		(0.16)	(0.21)		(0.13)	(0.31)	
	24	0.09	3.41	–	0.17	3.61	–	0.19	3.36	–	0.21	3.21	5.13	0.23	3.23	5.29
		(0.01)	(0.41)		(0.26)	(0.23)		(0.03)	(0.32)		(0.18)	(0.22)	(0.27)	(0.17)	(0.24)	(0.25)
	36	0.25	3.35	–	0.38	3.43	–	0.43	3.53	–	0.46	3.26	5.23	0.50	3.31	5.36
		(0.04)	(0.39)		(0.20)	(0.21)		(0.06)	(0.31)		(0.15)	(0.25)	(0.28)	(0.16)	(0.28)	(0.27)
	48	0.85	3.02	5.18	0.75	3.25	5.23	0.93	3.28	5.31	1.00	3.34	5.33	1.09	3.44	5.60
		(0.21)	(0.33)	(0.35)	(0.22)	(0.18)	(0.35)	(0.19)	(0.29)	(0.35)	(0.15)	(0.24)	(0.23)	(0.12)	(0.24)	(0.25)

Table 16: Means and st. dev. of modal estimates of ω_n^2 in DGP 2 (trimodal, $T = 8$)

Model		$N = 70$			$N = 210$			$N = 490$			$N = 980$			$N = 1,960$		
	k	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**	**Max1**	**Max2**	**Max3**
Real		1.36	3.68	5.82	1.36	3.68	5.82	1.36	3.68	5.82	1.36	3.68	5.82	1.36	3.68	5.82
MXL−N Pref.	6	3.86	–	–	3.97	–	–	4.04	–	–	4.08	–	–	4.15	–	–
		(0.45)			(0.46)			(0.44)			(0.43)			(0.44)		
MXL−N WTP	6	3.98	–	–	3.95	–	–	4.14	–	–	4.12	–	–	4.20	–	–
		(0.50)			(0.48)			(0.42)			(0.47)			(0.42)		
LML-Poly	12	0.31	3.33	–	0.43	3.31	–	0.76	3.56	–	0.81	3.60	–	0.91	3.68	–
		(0.06)	(0.40)		(0.05)	(0.28)		(0.13)	(0.29)		(0.08)	(0.30)		(0.14)	(0.23)	
	24	0.28	3.30	–	0.25	3.71	–	0.30	3.81	–	0.32	3.39	5.31	0.35	3.40	5.35
		(0.05)	(0.35)		(0.06)	(0.31)		(0.09)	(0.32)		(0.05)	(0.26)	(0.46)	(0.09)	(0.23)	(0.44)
	36	0.29	3.33	–	0.67	3.20	5.28	0.70	3.40	5.43	0.74	3.47	5.32	0.83	3.51	5.44
		(0.05)	(0.36)		(0.11)	(0.30)	(0.55)	(0.16)	(0.26)	(0.51)	(0.11)	(0.24)	(0.52)	(0.11)	(0.25)	(0.40)
	48	0.71	3.12	5.16	0.84	3.45	5.40	0.90	3.39	5.36	0.95	3.51	5.52	1.07	3.55	5.66
		(0.13)	(0.35)	(0.53)	(0.12)	(0.29)	(0.48)	(0.19)	(0.28)	(0.46)	(0.09)	(0.27)	(0.45)	(0.09)	(0.24)	(0.37)
LML-Step	12	0.38	3.43	–	0.39	3.50	–	0.49	3.71	–	0.52	3.79	–	0.58	3.94	–
		(0.05)	(0.38)		(0.06)	(0.30)		(0.05)	(0.32)		(0.05)	(0.26)		(0.08)	(0.25)	
	24	0.18	3.48	–	0.18	3.36	–	0.18	3.32	–	0.19	3.26	5.19	0.21	3.30	5.36
		(0.04)	(0.36)		(0.07)	(0.30)		(0.06)	(0.29)		(0.04)	(0.25)	(0.46)	(0.03)	(0.30)	(0.44)
	36	0.29	3.53	–	0.66	3.21	5.27	0.71	3.31	5.45	0.78	3.39	5.29	0.82	3.49	5.43
		(0.05)	(0.35)		(0.11)	(0.26)	(0.43)	(0.12)	(0.28)	(0.45)	(0.13)	(0.27)	(0.42)	(0.10)	(0.26)	(0.45)
	48	0.82	3.17	5.31	0.82	3.24	5.46	0.86	3.33	5.35	0.93	3.39	5.36	1.04	3.59	5.69
		(0.12)	(0.29)	(0.49)	(0.14)	(0.26)	(0.45)	(0.14)	(0.25)	(0.40)	(0.16)	(0.24)	(0.40)	(0.13)	(0.22)	(0.38)
LML-Spline	12	0.41	3.46	–	0.62	3.48	–	0.94	3.70	–	1.01	3.89	–	1.08	4.05	–
		(0.07)	(0.40)		(0.15)	(0.24)		(0.13)	(0.32)		(0.16)	(0.26)		(0.18)	(0.34)	
	24	0.09	3.43	–	0.17	3.78	–	0.20	3.40	–	0.21	3.30	5.20	0.24	3.35	5.48
		(0.03)	(0.34)		(0.04)	(0.28)		(0.09)	(0.32)		(0.05)	(0.26)	(0.25)	(0.05)	(0.27)	(0.32)
	36	0.34	3.51	–	0.71	3.31	5.38	0.44	3.32	5.40	0.47	3.27	5.45	0.52	3.41	5.46
		(0.07)	(0.33)		(0.12)	(0.26)	(0.41)	(0.10)	(0.30)	(0.38)	(0.07)	(0.26)	(0.25)	(0.14)	(0.28)	(0.35)
	48	0.85	3.14	5.41	0.92	3.34	5.32	0.93	3.41	5.43	1.02	3.42	5.59	1.15	3.60	5.63
		(0.14)	(0.36)	(0.35)	(0.14)	(0.28)	(0.35)	(0.12)	(0.32)	(0.35)	(0.13)	(0.28)	(0.31)	(0.16)	(0.28)	(0.30)

Table 17: Information criteria for tap water models.

Model	κ	$\ln \mathscr{L}^*$	**AIC**	**BIC**
MXL-N Pref.	11	−2,932	5,821	5,823
MXL-N WTP	11	−2,908	5,794	5,771
LC 2 classes	23	−2,896	5,741	5,753
LC 3 classes	35	−2,877	5,724	5,745
LML-Poly	22	−2,818	5,614	5,637
LML-Poly	33	−2,774	5,526	5,549
LML-Poly	44	−2,732	5,442	5,465
LML-Poly	55	−2,718	5,414	5,437
LML-Step	22	−2,802	5,582	5,605
LML-Step	33	−2,758	5,494	5,517
LML-Step	44	−2,716	5,410	5,503
LML-Step	55	−2,702	5,382	5,505
LML-Spline	22	−2,786	5,550	5,573
LML-Spline	33	−2,742	5,462	5,485
LML-Spline	44	−2,700	5,378	5,401
LML-Spline	55	−2,686	5,350	5,412

Table 18: Modal values of distributions of attributes' coefficients (Empirical application)

		Odor			Taste			Turbidity			Stain
Model/Attribute	κ	**Weekly**	**Monthly**	**Never**	**Weekly**	**Monthly**	**Never**	**Mild**	**Medium**	**Extra**	**Present**
MXL-N Pref.	11	1	1	1	1	1	1	1	1	1	1
MXL-N WTP	11	1	1	1	1	1	1	1	1	1	1
LML-Poly	22	2	2	2	1	2	1	1	2	1	2
LML-Poly	33	1	2	2	1	2	2	2	2	1	2
LML-Poly	44	2	3	2	2	3	2	3	2	3	2
LML-Poly	55	2	3	2	2	3	2	3	2	3	2
LML-Step	22	1	2	1	1	2	2	2	2	1	2
LML-Step	33	2	2	2	2	2	1	1	2	1	2
LML-Step	44	2	3	3	2	2	2	3	2	3	2
LML-Step	55	2	2	3	2	3	2	3	2	3	2
LML-Spline	22	2	2	1	1	2	1	1	2	1	2
LML-Spline	33	2	2	2	2	2	1	2	2	2	2
LML-Spline	44	2	3	2	2	2	2	3	2	3	2
LML-Spline	55	2	3	2	2	2	2	3	3	3	2

References

Abe, M. (1999), 'A generalized additive model for discrete-choice data', *Journal of Business & Economic Statistics* **17**(3), 271–284.

Bajari, P., Fox, J. and Ryan, S. P. (2007), 'Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients', *The American Economic Review* **97**(2), 459–463.

Balcombe, K. and Fraser, I. (2011), 'A general treatment of "don't know" responses from choice experiments', *European Review of Agricultural Economics* **38**(2), 171–191.

Bansal, P., Daziano, R. and Achtnicht, M. (2016), Comparison of parametric and semi-nonparametric representations of unobserved taste heterogeneity in discrete choice, Technical report, Social Science Research Network.

Bartczak, A., Chilton, S., Czajkowski, M. and Meyerhoff, J. (2017), 'Gain and loss of money in a choice experiment. The impact of financial loss aversion and risk preferences on willingness to pay to avoid renewable energy externalities.', **65**, 326âĂŞ334.

Bastin, F., Cirillo, C. and Toint, P. L. (2010), 'Estimating nonparametric random utility models with an application to the value of time in heterogeneous populations', *Transportation Science* **44**(4), 537–549.

Bertram, C., Meyerhoff, J., Rehdanz, K. and Wüstemann, H. (2017), 'Differences in the recreational value of urban parks between weekdays and weekends: A discrete choice analysis.', *Landscape and Urban Planning* **159**, 5–14.

Boeri, M. and Longo, A. (2017), 'The importance of regret minimization in the choice for renewable energy programmes: Evidence from a discrete choice experiment.', *Energy Economics* **63**, 253–260.

Boyd, J. H. and Mellman, R. (1980), 'The effect of fuel economy standards on the us automotive market: an hedonic demand analysis', *Transportation Research Part A: General* **14**(5-6), 367–378.

Brouwer, R., Martin-Ortega, J., Dekker, T., Sardonini, L., Andreu, J., Kontogianni, A., Skourtos, M., Raggi, M., Viaggi, D., Pulido-Velazquez, M., Rolfe, J. and Windle, J. (2015), 'Improving value transfer through socio-economic adjustments in a multicountry choice experiment of water conservation alternatives', **59**, 458–478.

Campbell, D. and Doherty, E. (2012), 'Combining discrete and continuous mixing distributions to identify niche markets for food', *European Review of Agricultural Economics* pp. 1–25.

Chaikaew, P., Hodges, A. and Grunwald, S. (2017), 'Estimating the value of ecosystem services in a mixed-use watershed: A choice experiment approach', *Ecosystem Services* **23**, 228–237.

Cherchi, E. and Polak, J. W. (2005), 'The assessment of user benefits using discrete choice models: implications of specification errors under random taste heterogeneity', *Transportation Research Record* **1926**, 61–69.

ChoiceMetrics (2009), 'Ngene 1.0 user manual and reference guide'.

Daly, A., Hess, S. and Train, K. (2012), 'Assuring finite moments for willingness to pay in random coefficients models', *Transportation* **39**(1), 19–31.

de Ayala, A., Hoyos, D. and Mariel, P. (2015), 'Suitability of discrete choice experiments for landscape management under the european landscape convention', *Journal of Forest Economics* **21**(2), 79–96.

Eilers, P. and Marx, B. (1996), 'Flexible smoothing with b-splines and penalties', *Statistical*

Science **11**(2), 89–121.

Ferrini, S. and Scarpa, R. (2007), 'Designs with a-priori information for nonmarket valuation with choice-experiments: a Monte Carlo study', *Journal of Environmental Economics and Management* **53**(3), 342–363.

Fosgerau, M. and Bierlaire, M. (2007), 'A practical test for the choice of mixing distribution in discrete choice models', *Transportation Research Part B* **41**(7), 784–794.

Fosgerau, M. and Hess, S. (2007), 'A comparison of methods for representing random taste heterogeneity in discrete choice models', *European Transport* **42**, 1—25.

Fosgerau, M. and Mabit, S. L. (2013), 'Easy and flexible mixture distributions', *Economics Letters* **120**(2), 206–210.

Fox, J. T., Ryan, S. P. and Bajari, P. (2011), 'A simple estimator for the distribution of random coefficients', *Quantitative Economics* **2**(3), 381–418.

Fukuda, D. and Yai, T. (2011), 'Semiparametric specification of the utility function in a travel mode choice model', *Transportation* **37**, 221–238.

Gouríeroux, C. and Monfont, A. (1996), *Simulation-Based Econometric Methods*, Oxford University Press.

Hanley, N., Colombo, S., Tinch, D., Black, A. and Aftab, A. (2006), 'Estimating the benefits of water quality improvements under the water framework directive: Are benefits transferable?', *European Review of Agricultural Economics* **33**(3), 391–413.

Hensher, D. and Greene, W. H. (2003), 'The Mixed Logit Model: The State of Practice', *Transportation* **30**(2), 133–176.

Hensher, D., Shore, N. and Train, K. (2005), 'Households' willingness to pay for water service attributes', *Environmental and Resource economics* **32**, 509–531.

Jacobsen, J. and Thorsen, B. (2010), 'Preferences for site and environmental functions when selecting forthcoming national parks', **69**(7), 1532–1544.

Johnston, R., Schultz, E., Segerson, K., Besedin, E. Y. and Ramachandran, M. (2013), 'Stated preferences for intermediate versus final ecosystem services: Disentangling willingness to pay for omitted outcomes', **42**(1), 98–118.

Juutinen, A., Kosenius, A. and Ovaskainen, V. . (2014), 'Estimating the benefits of recreation-oriented management in state-owned commercial forests in Finland: A choice experiment', *Journal of Forest Economics* **20**(4), 396–412.

Kneib, T., Baumgartner, B. and Steiner, W. (2007), 'Semiparametric multinomial logit-models for analysing consumer choice behaviour', *Advances in Statistical Analysis* **91**, 225–244.

Li, B. (2011), 'The multinomial logit model revisited: A semi-parametric approach in discrete choice analysis', *Transportation Research Part B: Methodological* **45**(3), 461–473.

McFadden, D. (1973), Conditional logit analysis of qualitative choice behavior, *in* P. Zarembka, ed., 'Frontiers in Econometrics', Academic Press, pp. 105–142.

McFadden, D. and Train (2000), 'Mixed MNL models for discete response', *Journal of Applied Econometrics* **15**, 447–470.

Mejía, C. and Brandt, S. (2015), 'Managing tourism in the Galapagos Islands through price incentives: A choice experiment approach', *Ecological Economics* (C), 1–11.

Mitchell, R. and Carson, R. (1989), *Using surveys to value public goods : the contingent valuation method*, Resources for the Future, Washington, D.C.

Morey, E. and Thiene, M. (2017), 'Can personality traits explain where and with whom you recreate? A latent-class site-choice model informed by estimates from a mixed-mode LC cluster models with latent-personality traits', *Ecological Economics* **138**, 223–237.

Nordén, A., J. Jönsson, A., Lagergren, F. and Lehsten, V. (2017), 'Divergence in stakeholders' preferences: Evidence from a choice experiment on forest landscapes preferences in Sweden', *Ecological Economics* **132**, 179–195.

Ohdoko, T. and Yoshida, K. (2012), 'Public preferences for forest ecosystem management in Japan with emphasis on species diversity', *Environmental Economics and Policy Studies* **14**(2), 147–169.

Revelt, D. and Train, K. (1998), 'Mixed logit with repeated choices: Households' choices of appliance efficiency level', *Review of Economics and Statistics* **80**(4), 647–657.

Rigby, D., Alcon, F. and Burton, M. (2010), 'Supply uncertainty and the economic value of irrigation water', *European Review of Agricultural Economics* **37**(1), 97–117.

Rose, J. M. and Bliemer, M. C. J. (2009), 'Constructing efficient stated choice experimental designs', *Transport Reviews* **29**, 587–617.

Rungie, C., Scarpa, R. and Thiene, M. (2014), 'The influence of individuals in forming collective household preferences for water quality', *Journal of Environmental Economics and Management* **68**, 161–174.

Sándor, Z. and Wedel, M. (2001), 'Designing conjoint choice experiments using managers' prior beliefs', *Journal of Marketing Research* **38**, 430–444.

Scarpa, R., Gilbride, T., Campbell, D. and Hensher, D. A. (2009), 'Modelling attribute non-attendance in choice experiments for rural landscape valuation', *European Review of Agricultural Economics* **36**(2), 151–174.

Scarpa, R., Thiene, M. and Marangon, F. (2008), 'Using flexible taste distributions to value collective reputation for environmentally friendly production methods', *Canadian Journal of Agricultural Economics* **56**, 145–162.

Scarpa, R., Thiene, M. and Train, K. (2008), 'Utility in WTP space: A tool to address confounding random scale effects in destination choice to the Alps', *American Journal of Agricultural Economics* **90**, 994–1010.

Scarpa, R. and Willis, K. G. (2010), 'Willingness-to-pay for renewable energy: Primary and discretionary choice of British households for micro-generation technologies', *Energy Economics* **32**(1), 129–136.

Scarpa, R., Willis, K. G. and Acutt, M. (2007), 'Valuing externalities from water supply: Status quo, choice complexity and individual random effects in panel kernel logit analysis of choice experiments', *Journal of Environmental Planning and Management* **50**, 449–466.

Thiene, M., Meyerhoff, J. and De Salvo, M. (2012), 'Scale and taste heterogeneity for forest biodiversity: Models of serial nonparticipation and their effects', *Journal of Forest Economics* **18**(4), 355–369.

Thiene, M. and Scarpa, R. (2009), 'Deriving and testing efficient estimates of WTP distributions in destination choice models', **44**(3), 379–395.

Thiene, M., Scarpa, R. and Louviere, J. (2015), 'Addressing preference heterogeneity, multiple scales and attribute attendance with a correlated finite mixing model of tap water choice', *Environmental and Resource Economics* pp. 1–20.

Train, K. (1998), 'Recreation demand models with taste differences over people', *Land Economics* **74**(2), pp. 230–239.

Train, K. (2008), 'EM algorithms for nonparametric estimation of mixing distributions', *Journal of Choice Modelling* **1(1)**, 40–69.

Train, K. (2009), *Discrete Choice Methods with Simulation*, second edition edn, Cambridge University Press, New York.

Train, K. E. (2016), 'Mixed logit with a flexible mixing distribution', *Journal of Choice Modelling* **19**, 40–53.

Train, K. and Sonnier, G. (2005), Mixed logit with bounded distributions of correlated partworths, *in* R. Scarpa and A. Alberini, eds, 'Applications of simulation methods in environmental and resource economics', Springer Publisher, Dordrecht, The Netherlands, chapter 7, pp. 117–134.

UNCI (2010), 'Data provided by unci unione nazionale industria conciaria national tannery industry association'.

Wedel, M., Kamakura, W., Arora, N., Bemmaor, A., Jeong, C., Elrod, T., Johnson, R., Lenk, P., Neslin, S. and Poulsen, C. S. (1999), 'Discrete and continuous representations of unobserved heterogeneity in choice modeling.', *Marketing Letters* **10**(3), 219 – 232.

Willis, K. and Scarpa, R. (2002), Yorkshire Water customers' utility and willingness-to-pay for changes in water service factor levels: A stated choice analysis. Report to Yorkshire Waters.

Willis, K., Scarpa, R. and Acutt, M. (2005), 'Assessing water company customer preferences and willingness-to-pay for service improvements: a stated choice analysis', *Water Resources Research* **41**, W02019.

Willis, K., Scarpa, R., Gilroy, R. and Hamza, N. (2011), 'Renewable energy adoption in an ageing population: Heterogeneity in preferences for micro-generation technology adoption', *Energy Policy* **39**(10), 6021–6029.

Yamamoto, Y. (2015), 'Opinion leadership and willingness to pay for residential photovoltaic systems', *Energy Policy* **83**, 185–192.

Yoo, J. and Ready, R. (2014), 'Preference heterogeneity for renewable energy technology', *Energy Economics* **42**, 101–114.