

UNIVERSITY OF WAIKATO

Hamilton

New Zealand

**Are Preferences for Food Quality Attributes
Really Normally Distributed?
An Analysis Using Flexible Mixing Distributions**

Vincenzina Caputo, Riccardo Scarpa,
Rodolfo M. Nayga and David L. Ortega

Working Paper in Economics 27/17

November 2017

Corresponding Author

Vincenzina Caputo

Department of Agriculture, Food
and Resource Economics
Michigan State University
Michigan
USA.

Phone: +1 517-884-8656

Email: vcaputo@msu.edu

Riccardo Scarpa

University of Waikato, NZ
University of Durham, UK
University of Verona, Italy

Email: riccardo.scarpa@univr.it

Rodolfo M. Nayga

Korea University
Norwegian Agricultural Economics
Research Institute
University of Arkansas
Arkansas

Email: rnayga@uark.edu

David L. Ortega

Department of Agricultural, Food
and Resource Economics
Michigan State University
Michigan
USA

Email: dlortega@msu.edu

Abstract

We empirically question the commonly employed distributional assumption of normality of taste distribution in mixed logit models with continuous random parameters. We use a WTP-space random utility discrete choice model with flexible distributions (Train 2016) on data from two choice experiments regarding beef with nested set of quality attributes. We specifically address distributional features such as asymmetry, multi-modality and range of variation, and find little support for normality. Our results are robust to attribute dimensionality in experimental design. Implications of our results for practitioners in the field are discussed.

Keywords

flexible taste distributions

mixed logit

logit mixed logit

food preferences

preference heterogeneity

1. Introduction

Product differentiation is a strategic tool for food market operators. Success in this area is heavily reliant on having adequate market information coming from reliable methods for differentiated consumer preference analysis. As a consequence, the mixed logit models for food choice data analysis introduced by Revelt and Train in 1998 were enthusiastically embraced by empirical researchers (Bonnet and Simioni 2001; Cicia, Del Giudice and Scarpa 2002; Lusk and Schroeder 2004; Alfnes *et al.* 2006; Rigby and Burton 2006) and are still widely used (Ortega *et al.* 2011; Caputo, Nayga and Scarpa 2013; Scarpa *et al.* 2013; van Wezemael *et al.* 2014; De Marchi *et al.* 2016; Bazzani *et al.* 2017). Operationalizing mixed logit models, however, requires assumptions on mixing preference distributions for the target population.

The question of what statistical distribution should be selected to model random taste coefficients to avoid unwarranted (and sometimes unintended) impacts in terms of data fit and welfare estimates, still poses serious empirical challenges to analysts. Like others before us, we start by observing that the assumptions on which these models are predicated, despite being often strong and crucial to the conclusions, are most often left unpersuasively justified. The contribution of this article is to explore the effectiveness of recently introduced tools for a robust investigation of such assumptions. Specifically, we offer some significant results on range, asymmetry and multimodality of taste distributions, which we deem as substantive for the future practice of food choice analyses. Our results also have significant implications for conceptual models of consumer demand whose results may be invalidated given their reliance on the assumption of uniform preferences (for example, Crespi and Marette 2003; Lapan and Moschini 2007; Giannakas and Yiannaka, 2008).

The use of various types of preference mixing—finite, continuous or a combination thereof—is by now the presumptive approach in the field of food choice, and it has been in many other areas of application (for example, environmental, health or transport economics). Yet, most published studies fail to explicitly report investigations on the sensitivity of their results to the sometimes crucial distributional assumptions under which they are derived. Furthermore, such assumptions are often predicated on weak arguments and motivation including operational convenience (for example, such as mathematical tractability), and comparisons of fit with alternative distributional assumptions. In this context, it is worth highlighting that consistency of maximum likelihood estimates holds only under the correct specification, and applies only probabilistically to the ‘comparatively’ best specification, especially when all the elements in the set of comparison share some shortcomings (for example, all imply symmetry to the mean).

Almost universally in our review of food choice applications, when the selected model allows for continuous mixing, it relies on parametric distributions (normal, log-normal, triangular, uniform, etc.). This approach is attractive because it reduces the space of parameters needed for model fit (for example, from quantiles to only first and second central moments),

but it overly simplifies matters, thereby ruling out several behaviourally plausible features of taste distributions, such as limited range, asymmetry, strong skewness and multimodality. This leads to inadequate conclusions, that often fit oddly in the face of common sense or even of mere introspection. Such discomfort has been expressed several times before and traces of it can be found in the concluding remarks of several previous papers approaching the issue from various perspectives (Train and Sonnier 2005, Cherchi and Pollak 2005, Burton, Balcombe and Rigby 2009). Warnings of significant biases due to erroneous distributional assumptions have been issued since the adoption of the mixed logit methodology. Yet, the issue has continued to receive little, if any, attention in empirical analyses of food choice.

To move the field forward, we explore the use of more robust approaches that can enable analysts to openly explore behaviourally realistic distributional structures of food taste. In practice, this requires the adoption of flexible distributional forms, such as mixture of parametric, semi-parametric or non-parametric approaches. There is some obvious resistance to adopting these approaches, as they are bound to be somewhat more complex to implement and they require larger sample sizes to achieve similar degrees of parameter estimates precision. Thus, a successful solution needs to be sufficiently practical to have wide applicability. In moving from a standard parametric description of preference variation to a more flexible one, the analyst faces several unfamiliar challenges linked to taste distributions. In this article, we focus on three important distribution features: the definition of the range of variation, symmetry and multi-modality. These features have obvious and important repercussions for the computation of statistical expectations and quantiles, which are crucial statistics in policy decisions. An example is the well-known so-called ‘fat-tail’ problem (for a recent review see Parsons and Myers 2016).

Throughout the article, we use a recently proposed semi-parametric choice model: the Logit-Mixed Logit (LML) developed by Train (2016) to explore the sensitivity of our results to the three distributional features mentioned above. This model allows for extremely flexible mixing distributions, that can accommodate asymmetry and multimodality, but it requires setting the range of variation. Hence, we also explore the stability of results in distributional outcomes by varying the range (the empirical support of the distribution). In addition, in response to recent works on the effect of choice context (Gao and Schroeder 2009; and Caputo, Scarpa and Nayga 2017), we also explore the sensitivity of our distributional results across food attribute types (for example, cue and independent) when increasing the number of attributes (from three to five) in the discrete choice experiment design and associated utility functions. Finally, to make the article more salient to recent tendencies in food choice, we decided to focus on random utility models specified in the WTP-space, so as to avoid scale issues and focus on value distributions.

This study contributes to the existing literature of consumer food preference analysis in three important ways. First, all food choice studies that addressed taste heterogeneity using continuous mixing have used parametric mixing distributions (that is, largely normal

distributions). Hence, they fail to simultaneously address the three issues we focus upon in this study. Two of these issues (multimodality and asymmetry) were addressed in Scarpa, Thiene and Marangon (2008), but they only applied a flexible semi-parametric distribution to one of the various random coefficients in their specification and they used a model in preference space. The present study is the first food choice application to simultaneously address the issues of range, symmetry and multi-modality for all random coefficients in WTP-space by using a flexible semi-parametric distribution. Our approach moves away from the standard assumptions of normality without excluding them. Our second contribution is to provide a specific exploration of estimates' sensitivity to the definition of the random coefficients' range of variation. Finally, to the best of our knowledge, this is the first study exploring the sensitivity of different distributional features across cue and independent attributes when extending the attribute space. As argued by Gao and Schroeder (2009) and Caputo, Scarpa and Nayga (2017), the way consumers value a 'cue' attribute (described as one whose levels correlate with the levels of other potentially absent attributes) and an independent attribute (relates to the physical aspects of the product whose information stands alone) can depend on the attribute space. Hence, this study adds to this stream of literature by showing that not only would consumers value these attributes differently across design dimensions, but also by suggesting that cue and independent attributes might be systematically characterized by different distributional features and context dependency.

The remainder of the article is articulated as follows. In the next section, we provide a brief and essential literature review as a background to highlight the glaring knowledge gap that we approach to fill. The third section provides a description of the data used. The fourth section discusses the method we employ and this is followed by a description of the estimation strategy and the discussion of the results. The final section presents our conclusions and some recommendations for changes in the practice.

2. Background

That the researcher's choice of taste distribution matters has been a central tenet of taste heterogeneity studies from its beginning. As early as 1999, Wedel *et al.* and later on in 2003 Hensher and Greene provided detailed guidance for its selection. A more recent review on the topic can be found in a working paper by Yuan, You, and Boyle (2015). Several early studies showed that parametric mixing distributions assumed ex-ante by researchers (for example, normal, lognormal, among others) may be limiting and may introduce mis-specification problems (Train and Sonnier 2005, Cherchi and Pollak 2005, Burton, Balcombe and Rigby 2009). These papers focused on bounding ranges of variation and therefore signs, and suggested remedies on how to handle distributions for theoretically signed coefficients (for example, for price) on the negative or positive orthants, and on asymmetry. For example, these papers explored parametric distributions or transformations thereof, which required further parameter estimates in the transformation function, often, as in the Johnson- S_B , of complex empirical identification. The evidence provided emphasised the vulnerability of results to bias

of different importance and size, in terms of post-estimation applications. Bias affects probability forecasts, marginal effects and welfare measures, all of which are of high relevance in food choice and food policy analysis.

Later studies have gone further in the direction of adding flexibility, often in an attempt to uncover multi-modality when present and of practical relevance. These studies have proposed either mixtures of parametric distributions (e.g. mixtures of normals Train 2008, Wasi and Carson 2013), or the use of either semi- or non-parametric mixing distributions (Bajari, Fox, and Ryan 2007; Fosgerau and Bierlaire 2007; Scarpa, Thiene and Marangon 2008; Train 2008; Bastin, Cirillo and Toint 2010; Fox, Ryan and Bajari 2011; and Fosgerau and Mabit 2013). Such distributions are more flexible in retrieving preference heterogeneity, thereby accommodating multimodality as well as asymmetry, and hence skewness. They may even come with the added bonus of being computationally less expensive in estimation (Train 2016; Bansal, Daziano, and Achtnicht 2016). However, because they are based on splines or polynomials, they are reliant on a larger parameter space than simply means and variances. Moreover, their sample-size requirements to achieve given degrees of accuracy are likely to be larger than those required by parametric distributions.

When the focus of taste heterogeneity is on economic values of food attributes, the typical subjects of investigation are distributions of marginal willingness to pay (mWTPs) or total welfare changes for selected food attributes. In linear utility specifications, these are non-linear functions of parameter estimates, such as ratios, and whenever price coefficients are random, the estimates of these functions are sensitive to distributional assumptions on the price coefficient. Early attempts to deal with this issue often resulted in studies in which the price coefficient was assumed to be fixed. This is, however, a scarcely defensible assumption, as it implies a fixed marginal utility of money. Other solutions rely on bounding its range of variation by, for example, using constrained triangular distributions (Alfnes *et al.* 2009; Hensher and Greene 2009; Scarpa *et al.* 2013; Hensher, Rose and Greene 2015) or the previously mentioned uniform or Johnson- S_B distributions.

A solution for this has been eloquently and persuasively discussed elsewhere (Train and Weeks 2005; Scarpa, Thiene, and Train 2008; Daly, Hess and Train 2012), and it suggests rescaling utility by the error scale. This solution was suggested earlier by Cameron and James (1987) in the context of contingent valuation, and it provides a specification of random utility directly in the WTP-space. Here, the random coefficients of attributes can be readily interpreted as marginal WTPs, and their distributions are derived in a manner less sensitive to the distributional assumptions for the price coefficient. However, up until now, they still have been reliant on parametric distributional assumptions (Balcombe, Burton, and Rigby 2011; Thiene, Scarpa and Marangon 2008).

Finally, Rose and Masiero (2010), argued that the assumptions implied by random utility models can be context dependent and affected by the nature of the datasets used and/or

dimensions of experimental designs. In food choice studies, for example, a number of recent papers have shown a specific interest in the sensitivity of marginal WTPs estimates to both expansion and hierarchy of food attributes (Gao and Schroeder 2009; Caputo, Scarpa and Nayga 2017). This literature explores the effects of progressively adding independent food attributes to choice contexts based on cue attributes in experimental choice. They found evidence of significant shifts in the means of the marginal WTPs, an issue which we also address in this study. As mentioned in the introduction, our study builds upon the issues discussed above and contributes to this body of extant literature.

3. Empirical Data

In our investigation, we use choice data from two choice experiments (A and B) exploring the effect of an incrementally larger set of attributes on beef selection. The dataset we use is part of a larger project investigating the effects of adding independent food attributes to cue attributes in discrete choice experiments published elsewhere (Caputo, Scarpa, and Nayga 2017). We use data from two experiments: Experiment A, which included only three beef attributes (*Certified U.S.*, *Guaranteed Tender*, and *Price*), and Experiment B, which added two more beef characteristics (*Guaranteed Lean*, *Sell-By Date*) for a total of five attributes. As in Caputo, Scarpa, and Nayga (2017), in this study we defined *Certified U.S.* as ‘cue attributes’, and *Guaranteed Tender*, *Guaranteed Lean*, and *Sell-By Date* as ‘independent attributes’. In both experiments the price attribute was specified with four levels: \$4.64; \$6.93; \$9.22; \$11.50. The other attributes were simply binary; they were either specified as present or absent in food product profiles. Each respondent was assigned to undertake a panel of eight choice tasks. Each task involved the selection of their preferred alternative out of three: two beefsteak profiles and the ‘no-purchase’ option. Sample statistics and further details about the experimental designs are reported in Caputo, Scarpa, and Nayga (2017). Table 1 shows the attributes and attribute levels included in this study and highlights the different use of the data from what done in Caputo, Scarpa, and Nayga (2017) and the present study.

4. Econometric Models

As previously discussed, we use a WTP-space utility specification in our analysis (Weeks and Train 2005). The objective of the investigation is to use flexible distributional assumptions for marginal WTPs, which allows us to retrieve more realistic taste distributions for food attributes because they allow for multimodality and asymmetry. We then contrast these flexible semi-parametric results with the standard parametric distributions based on normality, derived by using the familiar context of the mixed logit model (MXL) for panel data, as described in Revelt and Train (1998) (see also Train 2009).

Table 1: Attributes and Experiments

Attributes (attribute levels)	Caputo, Scarpa and Nayga 2017						Present Study	
	Experiment A		Experiment B		Experiment C		Experiment A	Experiment B
	A1	A2	B1	B2	C1	C2	From A1	From C1
Price (\$4.64;\$6.93; \$9.22; \$11.50)	√	√	√	√	√	√	√	√
Certified U.S. Product (absent/not absent)	√	√	√	√	√	√	√	√
Guaranteed Tender (absent/not absent)	√	√	√	√	√	√	√	√
Guaranteed Lean (absent/not absent)		√	√	√	√	√		√
Days before Sell-by Data (2 days; 8 days)				√	√	√		
Enhanced Omega-3 fatty acids (absent/not absent)						√		
Number of respondents	201		183		208		201	208

The flexible distribution approach is to be implemented by using the logit mixed logit (LML) model recently proposed by Train (2016). If the data display evidence of multimodality and asymmetry for some attribute mWTP, the flexible approach will make it apparent, while the MXL with normal distributions will hide this feature. For example, in Scarpa, Thiene and Marangon (2008), a random coefficient attribute that when assumed to be distributed normal showed an insignificant mean estimate with value close to zero and a very large standard deviation, once its distribution was evaluated semi-parametrically (using the Legendre polynomial method proposed by Fosgerau and Bierlaire 2007), it showed a much more plausible bi-modal distribution. The two modes, one at each side of zero made it clear that some consumer types desired the attribute and others objected to it. The normal interpretation, instead, implied indifference to the attribute, a difference with clear implications for marketing.

However, the investigation of the sensitivity of results to the range of variation, which needs to be defined a-priori for the LML, needs some decision rule. Train (2016) uses a range spanning two standard deviations (2SD) at both sides of the estimated mean. So, to start with, we adopt this approach too, which should work if the real range of variation is symmetric around the mean. Yet, in the presence of fat tails or multimodality, it might be inadequate. In this case, one can obtain guidance on how to extend the range to investigate by visual inspection of the histogram depicting mixing distributions resulting from the LML approach. More on this issue is reported in the estimation strategy section. We now proceed by briefly detailing the nature of both models, but we direct the readers interested in the details to the seminal papers.

Utility in WTP-Space

Following Train and Weeks (2005), the utility that individual n derives from choosing alternative j within a choice set J in choice situation t can be expressed as follows:

$$(1) \quad U_{njt} = V_{njt} + \varepsilon_{njt} = \tau_n(-Price_{jt} + \omega'_n \mathbf{x}_{jt}) + \varepsilon_{njt}$$

where V_{njt} is the observed portion of the utility; ε_{njt} is the error term, assumed to follow a Gumbel distribution; τ_n is a random positive scalar representing the price/scale parameter; here $Price_{jt}$ is the price level for 12 ounce of beef steak for alternative j and choice situation t , ω_n is a vector of estimated marginal WTPs; \mathbf{x}_{jt} is a vector of the observed non-price attributes for alternative j . In our application, these attributes are: US (*Certified US product*) and Tender (*Guaranteed Tender*) in experiment A, while in experiment B two more are added: Lean (*Guaranteed Lean*), and Sell (*Sell-by Date*).

Panel Mixed Logit

Let $y_{njt} = 1$ if individual n chooses alternative j in choice situation t , and 0 otherwise. Conditional on the vector $\langle \tau_n, \omega_n \rangle$, the probability of a sequence of T choices, assuming independence between choices is:

$$(2) \quad L_{njt}(\tau_n, \omega_n) = \prod_{t=1}^T \prod_{j=1}^J \left[\frac{\exp(\tau_n(-Price_{jt} + \omega'_n x_{jt}))}{\sum_{i \in J} \exp(\tau_n(-Price_{it} + \omega'_n x_{it}))} \right]^{y_{njt}}$$

The unconditional probability requires integrating over the distribution of the random parameter across respondents. To simplify notation, let us re-define $\langle \tau_n, \omega_n \rangle$ as β_n . The unconditional probability of the sequence of alternatives chosen by individual n can be expressed as follows:

$$(3) \quad P_n\{\Theta\} = \int L_{njt}(\beta_n) f(\beta_n|\Theta) d\beta_n$$

where $f(\beta_n|\Theta)$ is the probability density function of the vector of random parameters, as defined by the hyper-parameters Θ .

In what we take as the reference model, the mixture for the random parameters β_n is multivariate normal, so $\beta_n \sim N(\mu, \Omega)$ and $\Theta = \langle \mu, \Omega \rangle$. In other words, the hyper-parameters are the mean vector μ and the variance and covariance matrix Ω . Note here that for each random WTP the mean, median and mode all coincide, and the range with meaningful symmetric density around the means is a function of Ω . All these are undesirable restrictions that are usefully relaxed in the flexible specification that we now describe, proposed by Train (2016).

Panel Logit Mixed Logit (LML)

Unlike the MXL, in the LML model the joint mixing distribution of the random parameters ω_n is assumed 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. As shown in Train (2016), the joint probability mass function of random parameters $\beta_r \in S$ in the LML is represented by a logit formula:

$$(4) \quad \Pr(\beta_n = \beta_r) \equiv W(\beta_r|\alpha) = \frac{\exp(\alpha' z(\beta_r))}{\sum_{s \in S} \exp(\alpha' z(\beta_s))}$$

where α is a vector of probability mass parameters and $z(\beta_r)$ defines the shape of the mixing distribution. Substituting in equation (3), the unconditional probability $P_n(\alpha)$ of the sequence of choices of individual n is then:

$$(5) \quad P_n(\alpha) = \sum_{r \in S} L_{njt}(\beta_n) \left[\frac{\exp(\alpha' z(\beta_r))}{\sum_{s \in S} \exp(\alpha' z(\beta_s))} \right].$$

Note that the hyper-parameter is now the vector α and that the flexibility depends on the nature of the logit transformation of the z functions, to which we now turn.

The z Functions in the LML

Following Train (2016) three types of z functions are adopted here: orthogonal polynomials (for model LML-poly), grids (step-functions) (for model LML-step), and splines (for model LML-spline).

In his 2016 seminal article, Train starts from showing how normality can be approximated by specifying z as a second order polynomial in β_r . More flexibility in the shape of the distribution, allowing for asymmetry and multimodality, can be achieved by polynomials with higher order than second (in our LML-poly we use orders two, four, six, and eight), bearing in mind that the number of inflection points is equal to the polynomial order minus one. Of the various categories of polynomials available, orthogonal polynomials, such as Legendre polynomials (but also Hermite, Jacobi, Chebyshev, Bernstein polynomials), have the advantage of having uncorrelated terms. Correlation across β_r can be achieved by using cross-products of only first order terms, which greatly reduces the number of necessary parameters.

A second type of $z(\beta_r)$ is represented by a step function based on a grid over the parameter ranges (that is, the support). Consider partitioning the set S into G possibly overlapping subsets H_g . Consider the probability mass $W(\beta_r|\alpha)$ being the same for all points in a given subset, but different across subsets. In this case (LML-step) we have the following probability mass function:

$$(6) \quad \Pr(\beta_n = \beta_r) \equiv W(\beta_r|\alpha) = \frac{\exp(\sum_{m=1}^M \alpha_m(\beta_r \in T_m))}{\sum_{s \in S} \exp(\sum_{m=1}^M \alpha_m(\beta_s \in T_m))}$$

This set up generates a type of latent class at each point, except that the parameter values of each class are predefined, instead of being the outcome of an estimation, as in the case of a standard latent class model. In practice, a computational limitation of this approach is that with many attributes in the utility function the number of evaluations becomes quickly infeasible, even with rather broadly-spaced grids. In this study we use LML-step with four, six, eight and ten mass points.

Splines can also be used (in LML-spline) as they conform to the $\alpha'z(\beta_r)$ format required in (5). To illustrate, take an interval for a single parameter β that goes from start point β_1 and end point β_4 , with $\beta_4 < \beta_2$ and consider the two intermediate points (knots) β_2 and β_3 , with $\beta_2 < \beta_3$. Using $I(\cdot)$ as an adequate indicator function, this gives rise to the following four elements of the vector $z(\beta)$:

$$(7) \quad \begin{aligned} z_1(\beta) &= \left(1 - \frac{\beta - \beta_1}{\beta_2 - \beta_1}\right) I(\beta \leq \beta_2), \\ z_2(\beta) &= \left(\frac{\beta - \beta_1}{\beta_2 - \beta_1}\right) I(\beta \leq \beta_2) + \left(1 - \frac{\beta - \beta_2}{\beta_3 - \beta_2}\right) I(\beta_2 < \beta \leq \beta_3), \\ z_3(\beta) &= \left(\frac{\beta - \beta_2}{\beta_3 - \beta_2}\right) I(\beta_2 < \beta \leq \beta_3) + \left(1 - \frac{\beta - \beta_3}{\beta_4 - \beta_3}\right) I(\beta_3 < \beta), \end{aligned}$$

$$z_4(\beta) = \left(\frac{\beta - \beta_3}{\beta_4 - \beta_3} \right) I(\beta_3 < \beta),$$

The elements of the vector α requiring estimation in this case are only three, since the height of the spline is standardized to one (only relative height matters). Note that in (5) it is $\exp(\alpha'z(\beta_r))$ that defines the probability mass, and hence this non-linear transformation changes the spline shape, allowing flexibility. In this study, we use LML-spline with two, four, six and eight knots.

5. Model Estimation Strategy and Results

The data from Experiment A with two food attributes and Experiment B with four food attributes are used to estimate separate MXL models with normal mixing distributions for all mWTPs, i.e., $\omega_n \sim N(\mu, \Omega)$ and lognormal distribution for the scale/price coefficient factor. We termed these specifications as MXL-N and we use the results as reference points for comparisons with the flexible distribution model. In our specification search, we estimate a range of flexible distribution models, with different z functions and increasing number of parameters to explore the sensitivity of our results to increased flexibility. Specifically, four LML-polynomial (with order being four, six, and eight), four LML-step (with four, six, eight, and ten ‘steps’ or mass points), and four LML-spline (with knots being two, four, six, and eight) models¹ are estimated from data from each experiment. These flexible distribution models were estimated by using $[0, 2]$ as the range of variation for the price/scale coefficient. To explore the sensitivity to range, we investigate three different ranges for the mWTPs for food attributes. The extreme values of these ranges define the highest and the lowest marginal WTP values in the parameter space S and are constructed using the following three approaches:

- (1) Two standard deviations above and below the mean marginal WTPs obtained from the MXL-N model, this is the approach used in the seminal paper by Train (2016);
- (2) We then extend the range to cover three standard deviations above and below the mean of marginal WTPs obtained from the MXL-N model, to explore behavior in the tails; and
- (3) We extended the upper or lower range limits any time a sufficiently high probability mass was observed at the lowest and/or highest bin of the histogram. That is, whenever the tails of the distribution derived from (1) and (2) above had large mass. This assessment was made by visual inspection, but formal tests can be used.

The rationale for extending the range in these cases rests on our desire to investigate whether the high mass probability is due to an accumulation of consumers predicted to have mWTPs values at the upper end of the range, but who in reality have higher values and should

¹ For all models, during estimation the probability integral in equation (3) was approximated by using 2000 random draws for each person in the sample.

hence have probability mass located outside the investigated range. Alternatively, these mass points at high/low mWTP values could be confirmed to be accurate representations of preference densities. Some degree of asymmetry is to be expected in these distributions because of the very nature of the attributes; however, the MXL-N model forces symmetry around the mean/median/mode. After ascertaining the robustness of distributional findings in terms of range, asymmetry and multimodality, we assess their repercussion comparatively to the MXL-N results and across the two experiments with varying number of attributes.

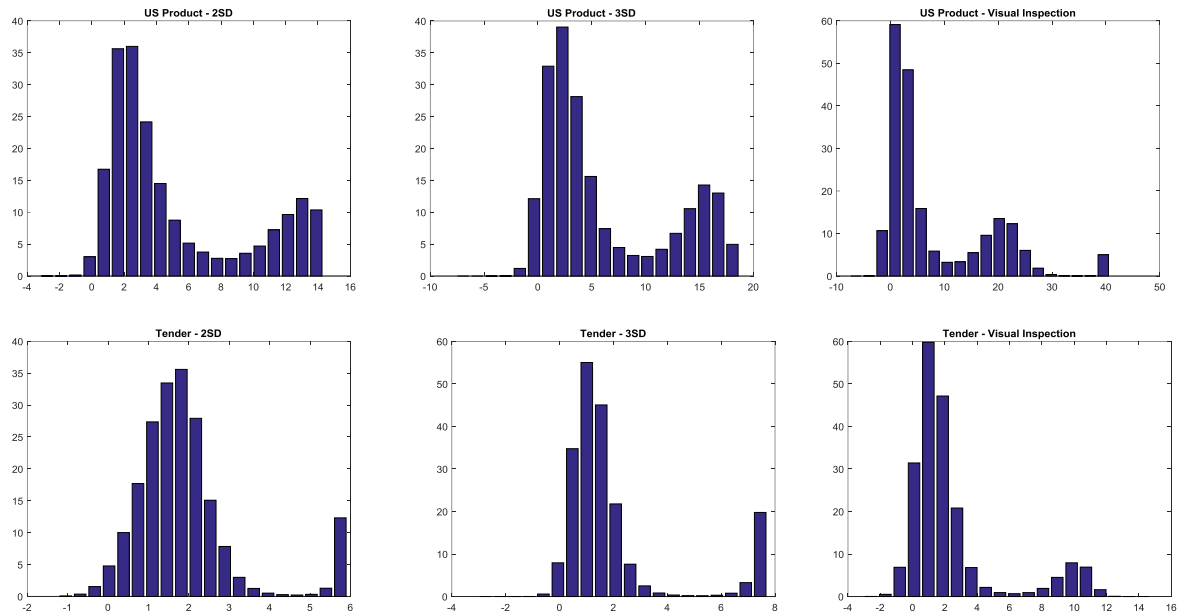
Data from each of the two experiments are used to estimate 24 models: four grid densities times three different ranges of variation times two experiments (A and B). This is repeated for each of the three types of z function (poly, step and spline), for a total of 52 flexible distribution models, respectively (26 per experiment).

The proper selection method for best performing models in the context of choice models with flexible semi-parametric distributions is still a subject of debate. In our case, we use standard information criteria that promote parsimony in the number of parameters: Akaike Information Criteria (AIC), the Bayesian Information Criteria (BIC), and modified Akaike Information Criteria (3AIC). The lower the information criterion value, the better the fit. Table 2 reports the model fit statistics for all models estimated across experiments A and B for each range approach utilized to define the highest and the lowest marginal WTP values in the parameter space S .

It can be noted that increasing the number of parameters improves the log-likelihood value, but does not necessary improve the information criteria values as these penalize for over-parameterization. This finding is consistent with Bansal, Daziano, and Achtnicht (2016), who employed the LML-polynomial, LML-step, and LML-spline models in both a Monte Carlo and empirical studies in the field of transportation. For ranges selected using the method of 2SD around the means estimated from the MXL-N, the best performing (accounting for all criteria) LML-polynomial models are those with fourth order polynomial in both experiments. In the LML-step models, it is with 6 steps and 4 for Experiment A and B, respectively, although for Experiment B the one with 8 steps has lowest AIC. For the LML-spline model, those with two knots outperform the rest in both experiments. Importantly, all flexible models outperform the MXL-N, except for the data in Experiment B but only when used in an extended asymmetric range. Exploring asymmetry seems to be more costly with over-parameterized models, which makes sense. For models with ranges established as 3SD around the MXL-N means, the best performing ones are those with the fewest parameters. This is true across all three z functions, although in Experiment B, the LML-step with 4 steps has better performance.

We now turn our attention to exploring the issue of asymmetry. To do so, we extend the range of variation for selected mWTPs based on visual inspection of the histogram representations of the mWTPs distributions from the 2SD and 3SD. These are reported in Figure 1 for the two steak attributes of Experiment A (US origin and certified tenderness).

**Figure 1: Estimated Distributions of Food Attribute Coefficients
Experiment A**



Both attributes show evidence of bi-modality in both 2SD and 3SD taste distributions, with high mass around small positive dollar values (0-6 for US and 0-3 for tender, both with highest mass at around 2 dollars), but also some high mass at the upper end of the dollar range. These upper tail values on the mWTP range are worth investigating further by extending the range. As a consequence, the upper limit in the third set of models for the tenderness attribute was extended from 6 and 8 dollars to 16, with the results of shifting and spreading the probability mass previously cumulated at 6 and 8 dollars over the range 8-12 dollars. A similar re-estimation for the US origin attribute, with range increased to a highly unlikely 50 dollars, shows that significant mass is still present at values over 20 dollars, with a third mode with mass at 40! This is brought about by a shift in the polynomial from the 4th to the 6th order. In fact, asymmetry in the range increases the number of parameters of the best fitting models across all z functions, for both experiments, except for Experiment B with LML-polynomial.

Table 2: Model Information Criteria of MXL-N and LML models, Experiments A and B

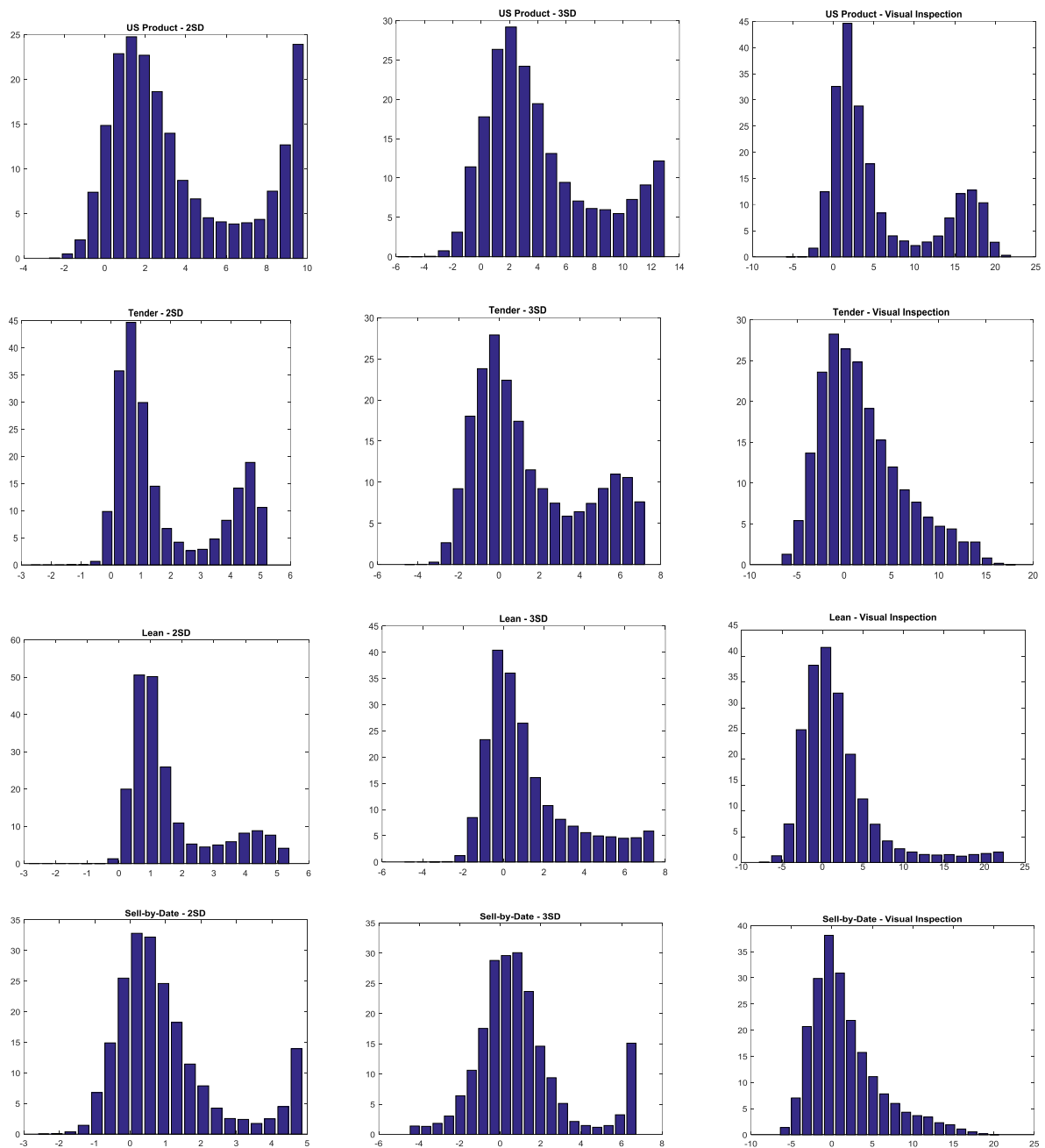
Model	LL	Par	BIC	AIC	3AIC	LL	Par	BIC	AIC	3AIC
	Experiment A (1,608 choices, N=201)					Experiment B (1,664 choices, N= 208)				
	Normal Distribution									
MXL-N	-1125	10	2324	2270	2280	-1294	21	2744	2630	2651
LML Polynomial	2SD above and below the mean									
4	-995	19	2129	2027	2046	-1239	34	2730	2546	2580
6	-986	27	2172	2027	2054	-1233	46	2807	2558	2604
8	-974	35	2206	2018	2053	-1225	58	2881	2566	2624
10	-974	43	2266	2035	2078	-1209	70	2937	2558	2628
LML Step										
4	-993	21	2142	2029	2050	-1248	38	2778	2573	2611
6	-979	29	2173	2016	2045	-1233	50	2836	2565	2615
8	-982	37	2237	2038	2075	-1220	62	2900	2564	2626
10	-968	45	2269	2026	2071	-1216	74	2981	2580	2654
LML Spline										
2	-987	21	2130	2017	2038	-1243	38	2768	2562	2600
4	-979	29	2173	2017	2046	-1230	50	2831	2560	2610
6	-974	37	2221	2022	2059	-1221	62	2902	2566	2628
8	-957	45	2247	2005	2050	-1210	74	2969	2568	2642
LML Polynomial	3SD above and below the mean									
4	-982	19	2104	2001	2020	-1265	34	2782	2598	2632
6	-977	27	2153	2007	2034	-1259	46	2859	2610	2656
8	-975	35	2209	2021	2056	-1245	58	2920	2606	2664
10	-972	43	2261	2029	2072	-1240	70	3000	2620	2690

Table 2 *continued*

Model	LL	Par	BIC	AIC	3AIC	LL	Par	BIC	AIC	3AIC
Experiment A (1,608 choices, N=201)						Experiment B (1,664 choices, N= 208)				
LML Step										
4	-987	21	2130	2017	2038	-1261	38	2805	2599	2637
6	-990	29	2194	2038	2067	-1261	50	2892	2622	2672
8	-981	37	2236	2037	2074	-1238	62	2936	2600	2662
10	-977	45	2286	2044	2089	-1236	74	3021	2620	2694
LML Spline										
2	-984	21	2122	2009	2030	-1270	38	2821	2615	2653
4	-978	29	2170	2014	2043	-1255	50	2882	2611	2661
6	-976	37	2225	2026	2063	-1245	62	2951	2615	2677
8	-969	45	2271	2028	2073	-1235	74	3019	2619	2693
LML Polynomial										
						Visual Inspection				
4	-1021	19	2182	2080	2099	-1353	34	2959	2775	2809
6	-1000	27	2200	2054	2081	-1349	46	3038	2789	2835
8	-994	35	2247	2059	2094	-1345	58	3121	2807	2865
10	-994	43	2306	2074	2117	-1341	70	3202	2822	2892
LML Step										
4	-1015	21	2186	2073	2094	-1367	38	3016	2810	2848
6	-990	29	2194	2038	2067	-1343	50	3056	2785	2835
8	-997	37	2268	2068	2105	-1331	62	3121	2786	2848
10	-993	45	2319	2077	2122	-1326	74	3200	2799	2873
LML Spline										
2	-1016	21	2187	2074	2095	-1365	38	3011	2806	2844
4	-993	29	2200	2044	2073	-1341	50	3053	2782	2832
6	-994	37	2261	2062	2099	-1331	62	3121	2786	2848
8	-981	45	2294	2052	2097	-1324	74	3197	2796	2870

We finally turn our attention to the stability of the distributional features to the addition of other food attributes in choice, by comparing the histograms for US origin and tenderness attributes of Experiment B (the two top rows of Figure 2) with 4 non-price attributes with the results obtained in Experiment A with only two (in Figure 1).

**Figure 2: Estimated Distributions of Food Attribute Coefficients
Experiment B**



Unexplained context-dependency of results is generally regarded as a negative feature in all methods, and this has been a criticism recently leveled to discrete choice models from experimental food data before (Gao and Schroeder 2009; Caputo, Scarpa, and Nayga 2017). This evidence, however, was obtained under parametric distributional assumptions. Whether this is still an issue with flexible functional forms is what we explore here. Comparing figures 1 and 2, we note that the bimodality of the taste distribution for the US origin and tenderness attributes are still supported by the results obtained with the symmetric ranges 2SD and 3SD. But once the asymmetric range is used, only the US attribute remains bimodal, while all other distributions appear unimodal and strongly skewed to the left—more so than what a normal distribution would correctly capture—and with well-behaved upper tails that taper out. Balcombe, Burton and Rigby (2009) already focussed on skewness and reported this to be an empirical regularity in preference distributions. Further, we note that the value range is less extended for these attributes in Experiment B than in Experiment A. This is consistent with what we expect in a choice context in which some cue attributes lose value in the presence of properly specified independent attributes, which would otherwise embed some value in the cue attributes when they are unspecified (Caputo, Scarpa and Nayga 2017). It is also confirmed by the mean and standard deviation values for the mWTPs reported in Table 3.

**Table 3: Statistics of Marginal WTP Estimates from MXL-N and LML
Bootstrapped Standard Errors, Models
Experiments A and B**

Models		Experiment A				Experiment B			
		MXL-N	LML-Polynomial		MXL-N	LML-Polynomial			
Variables	Par		2SD	3SD	Vis. Insp.		2SD	3SD	Vis. Insp.
<i>US</i>	μ	5.56* (0.47) ¹	5.13* (0.24)	6.14* (0.16)	7.92* (0.60)	3.53* (0.30)	3.75* (0.26)	4.29* (0.37)	5.87* (0.21)
	σ	4.39* (0.64)	4.32* (0.94)	5.76* (0.78)	9.36* (1.36)	3.17* (0.33)	3.36* (0.16)	3.94* (0.15)	6.46* (0.42)
<i>Tender</i>	μ	2.35* (0.37)	1.89* (0.35)	2.18* (0.31)	2.43** (1.30)	1.26* (0.34)	1.82* (0.14)	1.51* (0.33)	1.89** (1.03)
	σ	1.79* (0.47)	1.29* (0.14)	2.11* (0.23)	2.97* (1.00)	1.99* (0.34)	1.71* (0.22)	2.70* (0.25)	4.40* (0.52)
<i>Lean</i>	μ					1.26* (0.31)	1.62* (0.05)	1.22* (0.19)	1.85* (0.53)
	σ					2.07* (0.29)	1.36* (0.26)	2.19* (0.31)	4.85* (0.62)
<i>Sell</i>	μ					1.13* (0.24)	1.02* (0.25)	1.00* (0.29)	1.61** (0.81)
	σ					1.88* (0.31)	1.49* (0.06)	2.26* (0.30)	4.31* (0.55)

Note: Asterisk (*) and double asterisk (**) denote coefficients significant at 1% and 5% respectively.

This evidence corroborates the hypothesis that some sensitivity of results to the choice context is present, even when using flexible distributions. Yet, the main non-normal features of the distributions of tastes for cue attributes seem relatively stable to context. Interestingly, extending the range to the right, which allows for asymmetry, in the mWTP for the sell-by date attribute produces an upper tail that tapers out, rather than the bimodal distribution portrayed in the symmetric 2SD and 3SD results. Once again, behavior on the tails matters, and it is best captured by the asymmetric range, as the 2SD and 3SD representation still indicate bimodality for taste of this attribute. Altogether, these results suggest significant departures from the standard normality assumptions commonly invoked by food choice analysts in existing preference heterogeneity studies.

6. Robustness Check of Observable vs. Unobservable Sources of Heterogeneity

Differences in consumer preferences for food attributes can be explained by observable and/or unobservable sources of preference heterogeneity. Observable sources of preference heterogeneity such as demographics are those known by the researcher. They are commonly incorporated into discrete choice models through interactions with the experimentally designed levels of the attributes. The basic assumption of this modeling approach is that consumer preferences are heterogeneous due, at least in part, to differences in preferences across diverse socio-demographic groups. However, unobservable sources of preference heterogeneity may still remain even after such interactions are accounted for. These are unknown to the researcher and they are typically modeled by assuming random taste variation via the estimation of MXL models, where the distribution of random coefficients is intended to approximate unobserved sources of preference heterogeneity.

A natural question to ask in our study is whether the distributional features identified by the LML for each attribute of interest are due to observed and/or unobserved sources of preference heterogeneity. Taking advantage of the fact that we collected socio-demographic data during the CE surveys so as to profile our respondents, the samples from both experiments (A and B) were used to estimate MXL-N models in WTP space, accounting for observed sources of preference heterogeneity. We did so by interacting the experimentally designed attribute levels with the individual characteristics of our respondents. If these interactions are statistically insignificant, then we can conclude that observed individual characteristics fail to explain preference heterogeneity around the mean (Hensher, Rose and Greene 2015). This does not imply absence of preference heterogeneity around the mean, but simply that the socio-demographic characteristics of respondents fail to account for it. The results are presented in Table 4 for both experiment A and B.

Table 4: Statistics of Marginal WTP Estimates from MXL-N Model including Socio-Demographics, Experiments A and B

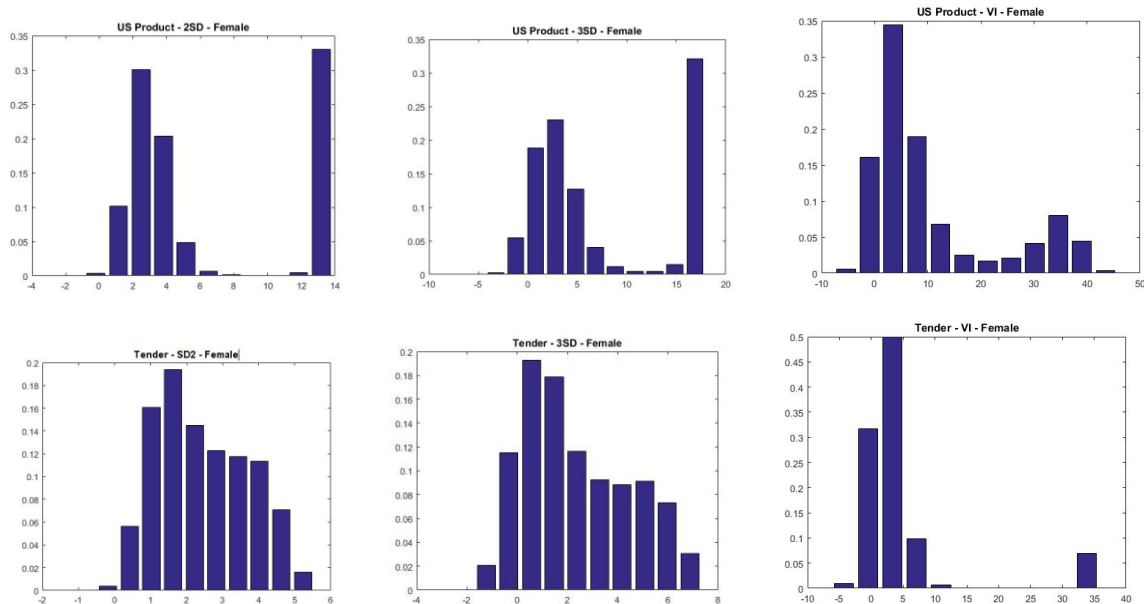
Models		Experiment A	Experiment B
Variables	Par	MXL-N	MXL-N
Main effects			
<i>US</i>	μ	5.22* (0.75)	3.05* (0.75)
	σ	4.02* (0.57)	3.13* (0.38)
<i>Tender</i>	μ	1.92* (0.67)	1.30* (0.58)
	σ	1.80* (0.44)	1.96* (0.40)
<i>Lean</i>	μ		1.34* (0.62)
	σ		2.06* (0.45)
<i>Sell</i>	μ		0.19*(0.51)
	σ		1.82*(0.29)
Interaction terms			
<i>US * Female</i>	μ	0.61** (0.31)	0.47 (0.29)
<i>US * Education</i>	μ	(0.068) (0.32)	0.16 (0.34)
<i>US * Age</i>	μ	(0.29) (0.33)	0.04 (0.30)
<i>US * Income</i>	μ	(0.28) (0.32)	(0.23) (0.30)
<i>Tender * Female</i>	μ	(0.08) (0.20)	(0.31) (0.29)
<i>Tender * Education</i>	μ	0.34 (0.230)	(0.10) (0.28)
<i>Tender * Age</i>	μ	(0.08) (0.23)	0.27 (0.26)
<i>Tender * Income</i>	μ	0.08 (0.23)	0.19 (0.28)
<i>Lean * Female</i>	μ		(0.29) (0.26)
<i>Lean * Education</i>	μ		(0.04) (0.26)
<i>Lean * Age</i>	μ		0.29 (0.25)
<i>Lean * Income</i>	μ		0.05 (0.26)
<i>Sell * Female</i>	μ		0.21 (0.23)
<i>Sell * Education</i>	μ		0.28 (0.23)
<i>Sell * Age</i>	μ		0.04 (0.22)
<i>Sell * Income</i>	μ		0.37 (0.24)
<hr/>			
Statistics			
<i>Choices</i>		1608	1664
<i>LL</i>		-1119.67	-1280.62
<i>Par</i>		18	37
<i>Number of Respondents</i>		201	208

Note: Asterisk (*) and double asterisk (**) denote coefficients significant at 1% and 5% respectively.

As can be seen from Table 4, none of the interaction terms between the demographic variables and the experimental design attributes are statistically significant across the two experiments, with the exception of gender in Experiment A. Hence, our findings generally confirm results from a number of applications of discrete choice models analyzing consumer food preferences, which have shown that demographic characteristics of respondents often fail to explain preference heterogeneity (Nilsson, Foster and Lusk 2006; Gracia, Loureiro and Nayga Jr. 2009; Caputo, Nayga and Scarpa 2013). Nilsson, Foster and Lusk (2006) suggest that the observable consumer characteristics might be poor indicators of food preference heterogeneity when analyzing consumer preferences for credence attributes of food products (for example, country of origin, brands, etc.) due to the strong reparability assumption between food attributes and demographic information.

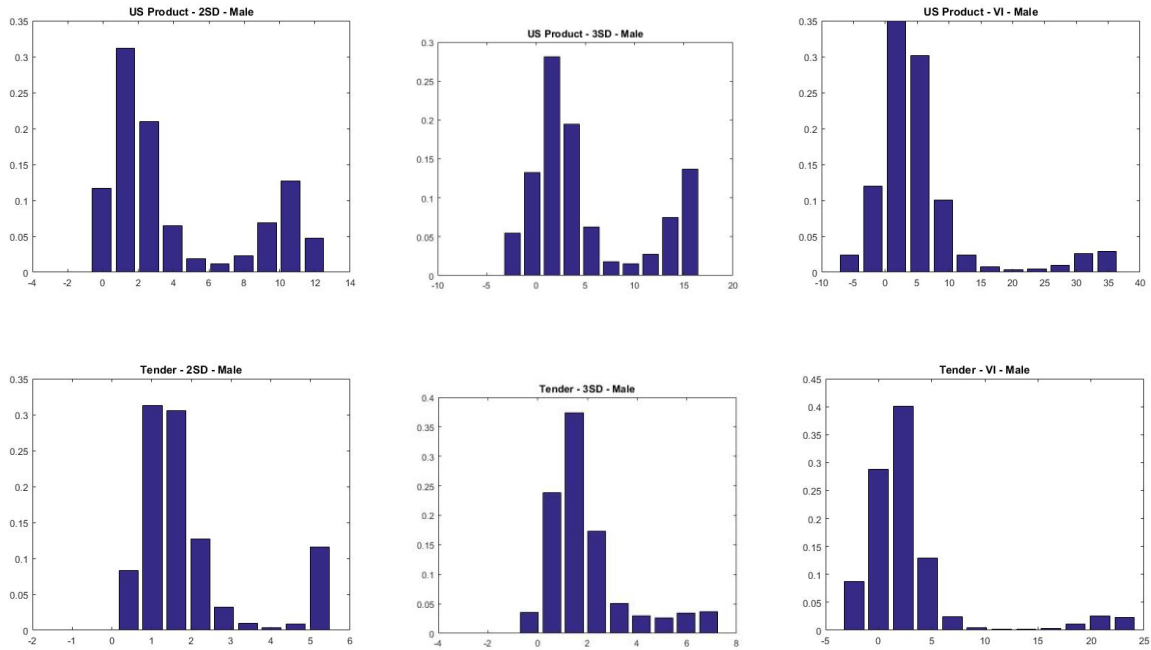
Given the significance of the interaction term between gender and the US Certified label in Experiment A, we estimated a LML² for each sub-sample based on gender (male and female) to further explore if there is heterogeneity in the estimates. As before, for each sub-sample, extreme marginal WTP values in the parameter space S are set to two and three standard deviations above and below estimated means of marginal WTPs from the MXL-N model with covariates. Also, any time a sufficiently high probability mass was observed at the lowest and/or highest bin of the histogram we extended the upper or lower range limits. Figures 3 (female sub-sample) and 4 (male sub-sample) report the estimated WTP distributions for *US certified* and for *guaranteed tender*.

Figure 3: Estimated Distributions of Food Attribute Coefficients for Female Experiment A



² Results of the LML by segmented samples (female and male) are reported in Appendix, Table A1.

**Figure 4: Estimated Distributions of Food Attribute Coefficients for Male
Experiment A**



Even after fitting different distributions by gender, in all cases a clear evidence of asymmetry—and to some extent of bimodality—remain. In other words, we still see departures from normality even in the sub-sample analysis by gender.

7. Conclusions and Recommendations for Practice

Food choice studies that addressed taste heterogeneity have used parametric mixing distributions (i.e., largely normal distributions) that fail to simultaneously address the three issues we focus upon in this study in relation to distribution features: the definition of the range of variation, symmetry and multi-modality. This is an important topic since these distribution issues could significantly affect the distribution of marginal WTP estimates that are used for important marketing and policy decisions. This study is the first to simultaneously focus on all these issues for all random coefficients of food attributes by using a flexible semi-parametric distribution estimated in WTP space.

Should future investigations of preference heterogeneity in food choice studies move beyond the parametric distributions in mixed logit? Our findings suggest that researchers using mixed logit models should check the robustness of their findings by also using flexible distributions. In our investigation on beef preferences, we use a flexible semi-parametric approach, the logit mixed logit estimator proposed by Train (2016) and discover that non-normal distributional features prevail. These features are sensitive to the setting of the range of variation and include acute skewness, asymmetry and bimodality. All of these features would

affect policy and marketing implications because they are likely to lead to different consumer stratifications from those derived using normality assumptions.

The approach used in this study is very flexible and not computationally burdensome, at least in our application. Our results suggest that the marginal WTP values show lower means in our experiments with a larger set of attributes, in accordance with previous findings in these contexts. Some significant probability mass extends over ranges of values that might appear very unlikely in reality, because they are excessively high. To limit this problem, and retain flexibility, we suggest that upper ranges for marginal WTP distributions from flexible distributions might need to be informed by responses to specific questions that can be included in survey questionnaires. In this way, the delimitation of the range could be grounded to some empirical data based, for example, on self-reported maximum willingness to pay statements for specific attributes.

To sum up, given our findings, future food choice analysts should consider systematic testing of the sensitivity of their results to the use of different parameter distributional features. Our hope is that this study will start a serious discussion about and consideration for this issue, given the increasing popularity of the use of discrete choice models in food choice studies. These studies are typically used not just for business applications, but also for welfare and policy analysis. Our results also have significant implications for research in other fields of inquiry where uniform type distributions between two extremes are commonly used (for example, studies investigating the market and welfare effects of novel food products and labels) since failure to capture deviations from normality could have serious economic consequences.

8. Reference list

- Alfnes, F., A. G. Guttormsen, G. Steine, and K. Kolstad. 2006. Consumers' Willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives. *American Journal of Agricultural Economics* 88: 1050–61.
- Bansal, P., R. A. Daziano, and M. Achtnicht. 2016. Comparison of Parametric and Semiparametric Representations of Unobserved Taste Heterogeneity in Discrete Choice. (December 28, 2016). Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2902402>.
- Bastin, F., C. Cirillo, and P. L. Toint. 2010. Estimating nonparametric random utility models with an application to the value of time in heterogeneous populations. *Transportation Science* 44(4): 537-549.
- Balcombe, K., M. Burton, and D. Rigby. 2011. Skew and attribute non-attendance within the bayesian mixed logit model. *Journal of Environmental Economics and Management* 62(3): 446-461.
- Bajari, P., J. T. Fox, and S. P. Ryan. 2007. Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients. *The American Economic Review*, 97(2): 459-463.
- Bazzani, C., V. Caputo, R. M. Nayga Jr., and M. Canavari. 2017. 'Testing Commitment Cost Theory in Choice Experiment.' *Economic Inquiry* 55(1): 383-396.
- Bonnet, C., and M. Simioni. 2001. Assessing Consumer Response to Protected Designation of Origin Labelling: A Mixed Multinomial Logit Approach, *European Review of Agricultural Economics* 28(2001): 433-449
- Burton, M., K. Balcombe, and D. Rigby. 2009. Mixed Logit Model Performance and Distributional Assumptions: Preferences and GM foods. *Environmental and Resource Economics* 42: 279-295.

- Cameron, T., and M. James. 1987. Efficient Estimation Methods for Closed-Ended Contingent Valuation Survey Data. *Review of Economics and Statistics* 69: 269-276.
- Caputo, V., R. M. Nayga, Jr., and R. Scarpa. 2013. Food miles or carbon emissions? Exploring labeling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural Economics* 57: 1-18.
- Caputo, V., R. Scarpa, and R. M. Nayga Jr. 2017. Cue versus independent food attributes: The Effect of Adding Attributes in Choice Experiments. *European Review of Agricultural Economics* 44(2): 211-230.
- Cherchi, E., and J. Polak. 2005. Assessing user benefits with discrete choice models: Implications of Specification errors under random taste heterogeneity. *Transportation Research Record: Journal of the Transportation Research Board* (1926): 61-69
- Cicia, G., T. Del Giudice, and R. Scarpa. 2002. Consumers' Perception of Quality in Organic Food: a random utility model under preference heterogeneity and choice correlation from rank-orderings. *British Food Journal* 104(3/4/5): 200-213.
- Crespi, J. M., and S. Marette, 2003. 'Does contain' vs. 'Does Not Contain': does it matter which GMO label is used?. *European Journal of Law and Economics* 16(3): 327-344.
- Daly, A.J., S. Hess, and K. E. Train. 2012. Assuring finite moments for willingness to pay in random coefficients models. *Transportation* 39(1): 19-31.
- De Marchi, E., V. Caputo, R. M. Nayga, and A. Banterle. 2016. Time preferences and food choices: Evidence from a choice experiment. *Food Policy* 62: 99-109.
- Fosgerau, M. and M. Bierlaire. 2007. A practical test for the choice of mixing distribution in discrete choice models. *Transportation Research Part B: Methodological* 41(7): 784-794.
- Fosgerau, M. and S. L. Mabit. 2013. Easy and flexible mixture distributions. *Economics Letters* 120(2): 206-210.
- Fox, J. T., S. P. Ryan, and P. Bajari. 2011. A simple estimator for the distribution of random coefficients. *Quantitative Economics* 2(3): 381-418.
- Gao, Z., and T. C. Schroeder. 2009. Effects of label information on Consumer willingness to pay. *American Journal of Agricultural Economics* 91(3): 795-809.
- Gracia, A., M. L. Loureiro, and R. M. Nayga Jr., 2009. Consumers' valuation of nutritional information: A choice experiment study. *Food Quality and Preference* 20: 463-471.
- Giannakas, K., and A. Yiannaka, 2008. Market and welfare effects of second-generation, consumer-oriented GM products. *American Journal of Agricultural Economics* 90(1): 152-171.
- Hensher, D. A., and W. H. Greene. 2003. The Mixed Logit Model: The State of Practice. *Transportation* 30(2): 133-76.
- Hensher, D. A. and W. H. Greene. 2009. Taming analytical distributions: valuation in WTP and utility space in the presence of taste and scale heterogeneity, Institute of Transport and Logistics Studies, University of Sydney, July.
- Hensher, D. A., J. M. Rose, and W. H. Greene, 2015. Applied Choice Analysis. Second Edition, Cambridge University Press.
- Lapan, H., and G. Moschini, 2007. Grading, minimum quality standards, and the labeling of genetically modified products. *American Journal of Agricultural Economics* 89(3): 769-783.
- Lusk, J. L., and T. C. Schroeder. 2004. Are Choice Experiments Incentive Compatible? A Test with Quality Differentiated Beef Steaks. *American Journal of Agricultural Economics* 86 (2): 467-82.
- Nilsson T., K. Foster, and J. L. Lusk, 2006. Marketing Opportunities for Certified Pork Chops. *Canadian Journal of Agricultural Economics* 54: 567-583.
- Ortega, D. L., H. H. Wang, L. Wu and N. Olynk. 2011. Modeling Heterogeneity in Consumer Preferences for Select Food Safety Attributes in China. *Food Policy* 36 (2): 318-324
- Parsons, G. R., and K. Myers. 2016. Fat tails and truncated bids in contingent valuation: An application to an endangered shorebird species. *Ecological Economics* 129: 210 - 219.
- Revelt, D., K. Train. 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics* 80(4): 647-657.
- Rigby, D. and Burton. 2006. Modelling Disinterest and Dislike: A Bounded Bayesian Mixed Logit Model of the UK Market for GM Food. *Environmental and Resource Economics* 33: 485-509.

- Rose, J., and L. Masiero. 2010. A comparison of the impacts of aspects of prospect theory on WTP/WTA estimated in preference and WTP/WTA space (Doctoral dissertation, Delft University of Technology).
- Scarpa, R., M. Thiene, and F. Marangon. 2008. Using flexible taste distributions to value collective reputation for environmentally friendly production methods. *Canadian Journal of Agricultural Economics* 56: 145-162.
- Scarpa, R., M. Thiene, and K. Train. 2008. Utility in Willingness to Pay Space: A tool to Address Confounding Random Scale Effects in Destination Choice to the Alp. *American Journal of Agricultural Economics* 90(4): 994-1010.
- Scarpa, R., R. Zanolli, V. Bruschi, and S. Naspetti. 2013. Inferred and Stated Attribute Non-Attendance in Food Choice Experiments. *American Journal of Agricultural Economics* 95 (1): 165-180.
- Train, K. E. 2008. EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling* 1(1): 40-69.
- Train, K. 2016. Mixed logit with a flexible mixing distribution. *Journal of Choice Modeling* 19-40-53.
- Train, K., and G. Sonnier. 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* 7: 117-134. In R. Scarpa and A. Alberini, eds.
- Train, K., and M. Weeks. 2005. Discrete choice models in preference space and willing-to-pay space. In R. Scarpa and A. Alberini, eds. *Applications of simulation methods in environmental and resource economics* 1: 1-16.
- van Wezemael, L. L., V. Caputo, R. M. Nayga Jr., G. Chryssochoidis, and W. Verbeke .2014. European consumer preferences for beef with nutrition and health claims: a multi-country investigation using discrete choice experiments. *Food Policy* 44: 167-176
- Wasi, N., and R. T. Carson. 2013. The influence of rebate programs on the demand for water heaters: The case of New South Wales. *Energy Economics* 40: 645-656.
- Wedel, M., W.A. Kamakura, N. Arora, A. Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin, and C.S. Poulsen. 1999. Discrete and continuous representations of unobserved heterogeneity in choice modeling. *Marketing Letters* 10: 219-232.
- Yuan, Y., W. You, and K. J. Boyle. 2015. A guide to heterogeneity features captured by parametric and nonparametric mixing distributions for the mixed logit mode. Selected Paper prepared for presentation at the Agricultural & Applied Economics Association and Western Agricultural Economics Association Joint Annual Meeting, San Francisco, CA.

Appendix A

**Table A1: Statistics of Marginal WTP
Estimates from the LML Model Across Demographics, Bootstrapped Standard Errors,
Experiment A**

Models		Female			Male		
		Polynomial			Polynomial		
		<i>2SD</i>	<i>3SD</i>	<i>Vis. Insp.</i>	<i>2SD</i>	<i>3SD</i>	<i>Vis. Insp.</i>
Variables	Par						
<i>US</i>	μ	6.49*	7.64*	10.67*	4.13*	5.15*	5.55*
		(0.35)	(0.45)	(2.34)	(0.18)	(0.74)	(0.59)
	σ	5.12*	7.16*	12.26*	3.91*	5.76*	8.16*
		(0.57)	(0.42)	(3.58)	(0.46)	(1.10)	(0.97)
<i>Tender</i>	μ	2.44*	2.37*	4.68*	1.94*	1.97*	2.87*
		(0.59)	(0.55)	(0.99)	(0.17)	(0.52)	(0.66)
	σ	1.26**	2.11*	8.63*	1.40*	1.62*	5.22*
		(0.67)	(0.33)	(3.34)	(0.31)	(0.42)	(2.32)
Statistics							
Choices		1056	1056	1056	552	552	552
LL		672.44	672.44	672.44	312.74	312.74	312.74
<i>Par</i>		19	19	19	19	19	19
<i>N of. Respondents</i>		132	132	132	69	69	69

Note

Asterisk (*) and double asterisk (**) denote coefficients significant at 1% and 5%, respectively.