Title

Normative Measurement of Consumer Preference Utilities

Abstract

There have largely been two main approaches for measuring consumer preferences, ad hoc conjoint analyses (CAs) or discrete choice experiments (DCEs). These approaches, while successful, also have drawbacks: CAs are limited by assumptions of additive attribute utilities, and are not grounded in formal models of human choice behavior. DCEs, while theoretically grounded, require careful experimental design, and preference utilities are obtained only by fitting the data with choice models, which also contain limiting assumptions. Thus, despite the abundance of preference measurement tools available, there is still a need for a normative, theoretically grounded method for measuring preferences directly. Here, we introduce such a method, discrete Markov Chain Monte Carlo with People (d-MCMP). Under basic assumptions about human choice behaviour, d-MCMCP is mathematically guaranteed to converge to the true distribution of utilities over full profiles, or a scaled version of this distribution. We show using both simulation and human experiment that d-MCMCP offers a powerful normative tool for measuring preferences that is particularly efficient when consumers are selective in their preferences.

Keywords: consumer preferences, normative tools, normative methods, measurement, consumer choice, choice utilities

Introduction

Consumer preference measurement has been widely and successfully applied to help market researchers with product development, pricing, segmentation, positioning, and advertising (Green and Krieger 2002; Green and Rao 1971; Green and Srinivasan 1978; Green and Srinivasan 1990; Hauser and Rao 2003; Johnson 1974; Louviere, Hensher, and Swait 2000; McFadden 1971; Orme 2009; Pekelman and Sen 1979; Rao 2008). There have largely been two general paradigms for evaluating consumer preferences. One tradition, broadly known as conjoint analysis (CAs) (Cattin and Wittink 1989) rises from the assumption that attribute utilities can be algebraically combined to predict the overall preference for a product. CAs may include the use of partial profile ratings (Bradlow, Hu, and Ho 2004; Green 1974), self explication methods (Park, Ding, and Rao 2008; Srinivasan 1988), and hybrid techniques, (Agarwal and Green 1991; Carmone 1987; Green 1984; Johnson 1987; Toubia, Hauser, and Garcia 2007). While widely used, CAs are usually constrained by limiting assumptions, such as strictly additive attribute utility functions (independence among part-worth attribute utilities), which can result in mischaracterizations of actual preference utilities. Additionally, CAs are primarily ad hoc approaches that are not theoretically grounded in formal models of human choice behaviour. Thus, their results cannot be analyzed in a way that is consistent with neoclassical economic theory, because ratings and attribute importance measures do not readily translate into choice or matching. Thus CA results are not appropriate for wider use in economic modelling and applied economics, such as welfare and policy assessment (Louviere, Flynn, and Carson 2010).

In a return to a more theoretically grounded, and ecologically natural approach, researchers have also developed choice-based methods known as discrete choice experiments (DCE) (Elrod, Louviere, and Davey 1992; Foxall 2007; Foxall 2010; Louviere 1988;

Louviere et al. 2000; McFadden 1971; Train 2003). In DCEs, respondents are asked to make a series of choices between two or more product profiles, thus using a task that captures the comparison based choice behaviour of natural shopping. DCEs are sometimes known as choice-based conjoint analysis, though the use of this terminology has also been cautioned because, in contrast to CAs, DCEs do have a solid theoretical basis in random utility theory (RUT) (see Louviere et al. for further discussion (2010)). RUT proposes that there exists a latent construct of "utility" for each choice alternative, and that these utilities consist of a systematic (explainable) component that is based on attributes that differ between the choice alternatives and a random (unexplainable) component. More formally this means that, for each individual, there exists a utility, $U_i=V_i+\varepsilon_i$, where V_i is the explainable component of choice, ε_i is the random component, and U_i is the utility. A variety of choice modelling procedures have been developed assuming that human choice behavior is described by a functional form based on these utilities (McFadden 1971; Train 2003). Full product utilities, as well as marginal, attribute utilities can then be inferred from these model fits.

While powerful and successful, DCEs do pose some limitations. One is that DCEs often require careful design to ensure the experiment is efficient. Much research has been devoted to the efficient design of choice experiments (Louviere et al. 2008; Louviere, Pihlens, and Carson 2011), with the more recent work often employing sophisticated Bayesian modelling (Johnson 1974; Kanninen 2002; Louviere et al. 2008; Louviere et al. 2011; Louviere and Woodworth 1983; Rossi, Allenby, and McCulloch 2005; Sándor and Wedel 2001; Sándor and Wedel 2005; Yu, Goos, and Vanderbroek 2009). However, these methods are complex to implement. Furthermore, in order to extract utilities (both partial utilities and product utilities), one has to fit a choice model to the data. Inherent in these choice models, are a priori assumptions about the structure of attribute utility dependencies (e.g. attributes are assumed to be either independent or contain 2^{nd} or correlations). Thus, the accuracy of the

utility values extracted from DCEs largely depend on whether the appropriate model has been chosen, the constraints imposed by the chosen model's parametric structure, and whether the data was sufficient for being fitted by the chosen model. Furthermore, there may be circumstances where a researcher wants to obtain the distribution of utility weights over complex products directly, without requiring the estimation of part-worth attribute utilities. Here, the complexity of experimental design and model fitting required of DCEs, seems perhaps needlessly cumbersome.

Thus, despite the abundance and variety of preference measurement methods available, the field of marketing is still in significant need of a simple, normative method that can directly measure preference utilities over complex products, and that is not constrained by assumptions about the form of attribute utility dependencies. The importance of normative, non-parametric approaches is manifold. First, non-normative approaches (that are not grounded in formal theories of choice behaviour) are not guaranteed to measure true preferences under most circumstances. Also, an approach grounded in formal theory allows for future developments and expansions on the approach that continue to be theoretically grounded. Furthermore, the results produced from normative tools will be consistent with economic demand theory, and thus can then be then reliably used in applied economics such as for welfare and policy assessments. The advantage of a non-parametric (no assumption about attribute utility dependencies) approach is that it allows the data to capture any form of nonlinear dependencies among attribute utilities. Though nature of these dependencies will still need to be obtained by fitting the data to models of how attribute utilities may combine (as with DCEs), the possible attribute interactions captured in the data are no longer limited by the a priori design of the experiment.

Here we present for the first time, such a normative, theoretically grounded method for measuring consumer preferences utilities over a wide array of complex products. The

procedure is known as discrete Markov Chain Monte Carlo with People (d-MCMCP) and uses a binary choice task on full product profiles. Based on broad, flexible assumptions of human binary choice behaviour, which are in accordance with RUT, d-MCMCP is mathematically guaranteed to converge to the true distribution of preference utilities, or a scaled version of this. d-MCMCP was recently proposed in the field of mathematical psychology as a method of investigating the structure of human categories (Hsu et al. n.d.; Martin, Griffiths, and Sanborn 2011; Sanborn, Griffiths, and Shiffrin 2010). However, the significance of d-MCMCP in relation to measuring preference utilities has not been recognized nor tested. In this work, we first demonstrate how d-MCMCP can be applied to measure consumer preferences. We then test its performance using simulations and experiment, and show that it is an especially effective method for measuring preferences among expert consumers who are selective in their preferences.

Deleted: 2011

The advantages that d-MCMCP offers are numerous: First the ability to directly capture individual preference distributions using full product profiles means it is easily applied to real products, i.e. the precise brand name and model can be used/included in the product profile. This is useful because there will be times where a researcher may wish to directly measure preferences over a range of real products, such as when predicting sales for a new stock, or in a new location. Second, the nonparametric nature of d-MCMCP means that results are not heavily dependent on the experimental design. No a priori assumptions need to be made about attribute utility dependencies. This allows the market researcher to obtain data that can be flexibly analyzed afterwards to test any number of hypotheses about preference structures with any type of dependencies among attribute utilities. For example, after conducting a d-MCMCP experiment, one could look for market segments of individual wine consumers who only want white wines if they are cheap and from Italy, but never would prefer cheap white wines otherwise. In contrast, in traditional CA or DCE approaches, this three-way interaction

between attributes would have to be hypothesized beforehand and incorporated into the experiment design. Finally, d-MCMCP is a highly efficient method, which capitalizes on the efficiency of a Markov Chain Monte Carlo algorithm to quickly explore preferences over a large range of profile options within a feasible number of binary choice trials. This efficiency arises because the algorithm adaptively focuses on offering choice profiles that will most effectively measure an individual respondent's preferences. Thus, individual preferences over a wide range of products can be obtained in a survey of feasible length, allowing for effective market segmentation analysis.

The outline of this paper is as follows. First we introduce the key ideas and methodological details behind the d-MCMCP method, and show how the method can be used to measure preferences. Next, we test the suitability of d-MCMCP for preference measurement by testing its convergence properties on simulated choice data in the example domain of wine preferences. Note that we have chosen the domain of wine preferences as it is a complex product of many attributes and levels. However, we would expect that the simulation results apply to any domain with a similar structure of attribute level utilities. In order to more clearly illustrate how d-MCMCP works, we will compare its performance to that of its naïve alternative, which is to estimate the distribution of profile utilities by constructing a histogram over random pair-wise choices. Finally, we then apply the d-MCMCP method in an online experiment to measure real human wine preferences. Again, we compare the d-MCMCP results with the random pair-wise choice method. We show that d-MCMCP is particularly useful for capturing profile preferences which arise from more selective attribute utilities, e.g. preferences for a select few, potentially correlated, attribute levels.

Discrete Markov Chain Monte Carlo with People

The d-MCMCP method is based on the Metropolis Hastings algorithm (Hastings 1970), a Markov Chain Monte Carlo (MCMC) sampling algorithm originally developed in statistical physics to sample from complex, intractable distributions. Recently, researchers in mathematical psychology have adapted MCMC sampling method so that is can be used with people to measure mental representations of categories and concepts (e.g. a dog, or a happy face) (Hsu et al. n.d.; Martin et al. 2011; Sanborn et al. 2010). These are known as MCMCP and d-MCMCP, which measure continuous and discrete sets of items respectively. The only difference between MCMCP and d-MCMCP is that d-MCMCP has been adapted to handle discrete sets of items. MCMCP and d-MCMCP measures the relative strengths to which items in a large choice set are representative of the concept in question (e.g. how strongly an image is indicative of a happy face). Here we observe that the discrete version, d-MCMCP, extends naturally to the measurement of people's preference utilities over a large choice set of product profiles. In contrast to the previous work in psychology, we now are measuring how strongly an item is representative of a preferred product rather than a category.

Formatted: German Germany

Field Code Changed

Deleted: 2011

Formatted: German Germany

Deleted: relative strengths of

Overview of a d-MCMCP experiment

The overview in this section applies equally to the continuous version, MCMCP, and the discrete version, d-MCMCP. However, d-MCMCP is more suitable to measuring preferences over real products, which are a discrete set of items, and which are usually not readily represented in a continuous space. Thus, we have used d-MCMCP in this work, and we will just make references to d-MCMCP here. In a d-MCMCP choice experiment, a human participant is asked to make pair-wise choices between items over many trials. During each trial, participants are presented with two items from the choice set. How these two items are chosen is a key feature of the method and will be explained below. Participants are then asked to choose one of the two items under a choice criterion. For the purposes of measuring

Deleted: our purposes

preferences, this criterion will be "which is the preferred product?" The chosen item in each trial is recorded, eventually resulting in a long list of choices made across the many trials. With enough trials, the relative frequency with which an item is chosen becomes proportional to the strength of preference for that item relative to others in the choice set.

A way to gain intuition for how d-MCMCP works is to first consider the more naïve approach of obtaining preferences through randomly presented pair-wise choices. In this approach, preferences can be obtained by showing many randomly chosen pairs of items and constructing a preference histogram over all chosen items. However, such an approach is limited because it is usually desirable to explore a choice set containing a large range of items, and it is not practical to show all possible pair-wise combinations, (e.g. to show every possible pair in a database of 2000 items would require about 2 million trials). d-MCMCP overcomes this problem by introducing an efficient way of exploring a large number of items: the algorithm focuses on accumulating choices over high preference regions and quickly moves away from regions of low preference. Additionally, choices over randomly chosen pairs of items is not mathematically guaranteed to converge to the correct distribution of actual relative preference strengths over the choice set, and indeed under most circumstances is unlikely to. In contrast, d-MCMCP, under simple general assumptions about human binary choice behaviour, have been mathematically shown to converge either to the true distribution over profile utilities, or a scaled version of this distribution (Sanborn et al. 2010).

d-MCMCP works by building a list of chosen items. For each trial there is a "current item", I_C, and a "proposal item", I_P. d-MCMCP starts with an initial "current item", I_C, typically chosen at random and a "proposal item" that is chosen to be "nearby" to the current item (e.g. a similar facial image). The human participant is then asked to choose whether I_C or I_P. is more preferred. The chosen item is recorded. If item I_P is chosen, it becomes the new current item (I_C). If I_c is chosen, it remains as current item for the next trial. A new trial

begins in which a new "nearby" proposal item is chosen based on the current item. The participant is again asked to choose between current and proposed items, the choice is recorded, and so forth.

Mathematical Details

Again, for the interest of clarity we will just refer to d-MCMCP, while most of the mathematical details described below apply equally to the continuous version, MCMCP. The only difference between MCMCP and d-MCMCP arises in the subsection *proposing nearby states*. This is because continuous MCMCP makes random Gaussian proposals on a continuous space, whereas d-MCMCP has been adapted to make proposals in a discrete space.

Metropolis MCMC Algorithm. d-MCMCP works based on the convergence properties of the Markov Chain Monte Carlo (MCMC) algorithm known as the Metropolis Hastings algorithm. MCMC algorithms are a class of methods for generating samples from complex probability distributions by constructing Markov chains that converge to those distributions over time (see (ref 9)). If we want to draw a sample from the probability distribution $p(x)$, we define a Markov chain such that the stationary distribution of that chain is $p(x)$, and sample a sequence of states from that chain. If the sequence is long enough, the states of the chain can be treated similarly to samples from $p(x)$. The Metropolis algorithm (Hastings 1970) is one of the most popular methods for constructing such a Markov chain. The sequence of states is initialized with an arbitrary value, x'. The next value in the sequence is generated via a two-step process. First, a candidate for the next value, x'', is chosen by sampling from an arbitrary proposal distribution conditioned on x' that is specified by the designer of the algorithm, $q(x'';x')$. Second, a decision is made as to whether that proposed value will be accepted using a valid acceptance function which is a function of the relative probability of x and x′ under

the target distribution *p(x)*. An example of a valid acceptance function is the Barker function (Barker 1965) which specifies the acceptance probability to be:

$$A(x^*; x) = \frac{p(x')}{p(x') + p(x'')} \quad (1)$$

and defines a Markov chain that converges to *p(x)* provided the proposal distribution *q(x′; x″)* is symmetric, with *q(x′; x″) = q(x″; x′).*

From MCMC to d-MCMCP. d-MCMCP method is based on a correspondence between human choice behaviour and the Barker acceptance function (equation 1). If a task can be constructed in which people are offered a choice between $x=x''$ and $x= x'$ and choose x' with probability $P_{choice}(x') = \frac{u(x')}{u(x'') + u(x')}$ (2)

then this provides a valid acceptance function for a Markov chain that will converge to *u(x),* up to a normalization constant. Thus, based on this correspondence between d-MCMCP and the Metropolis algorithm, people's choices can then be used to determine which proposals are accepted (Sanborn et al. 2010). In a standard experiment, people would be asked to make a series of pair-wise decisions in which they are asked to choose the best category member from two proposed stimuli. The stimuli that are presented in each decision correspond to the values x and x' in the Metropolis algorithm, and the choices that people make determine which proposals are accepted. With enough decisions, d-MCMCP will converge to samples from the probability distribution associated with that category, and individual stimuli will be encountered with probability given by *u(x)* (up to normalization constant). The proposal distribution can be selected by the experimenter, provided it is symmetric in the way required by the Barker acceptance rule. For the case of consumer preferences *u(x)* will be the utility for the item *x.*

Equation 2 has a long history as a model of human choice probabilities, where it is known as Luce's choice rule or the ratio rule (Luce 1963). This rule has been shown to provide a

good fit to human data when people choose between two stimuli based on a particular property (Bradley 1954; Clarke 1957; Hopkins 1954). Under more general assumptions for human binary choice behaviour, such as the logit assumption often used in consumer choice modelling, the Markov chain will converge to a scaled version of the distribution as long as people's choices scale monotonically with equation 2 (Sanborn et al. 2010). In particular if

$$P_{choice}(x') = \frac{f(u(x'))}{f(u(x')) + f(u(x''))}$$

then the choices are guaranteed to converge to be samples from the distribution *f(u(x))*. So for example, if choices were assumed to follow a binary logit model, with utilities $U'=A\,x'+e$ and $U=Ax+e$, where *e* is an independent and identically distributed extreme valued random noise term with zero mean, then $P_{choice}(x') = \frac{\exp(Ax')}{\exp(Ax') + \exp(A(x''))} + e$ and the choices are guaranteed to converge to the distribution *exp(u(x))* (the random noise will average out), in which case *u(x)* can be recovered by taking the log of the distribution. In general, if the shape of *f(x)* is unknown, as long as it's a monotonic function, the converged distribution can be taken to be a scaled version of the actual distribution.

Proposing nearby states. In order for d-MCMCP to most efficiently explore the space of items, one has to have an appropriate method of making a "nearby" proposal. A key assumption in using the Barker acceptance function is that the proposals must be symmetric. That is the probability of choosing a proposal value given a current value is the same if the proposal and current values were reversed. Thus, for each profile item, one must identify a set of "neighbors" in such a way that the neighbor relationship is symmetric, i.e. if x' is a neighbor of x'', then x'' is a neighbor of x'. To achieve this one does the following: First, apply a similarity measure to every pair of items in the choice set. Any reasonable measure of similarity can be used, for example the number of attribute levels in common between two

items. Next, we need to build an interconnected graph between items in the choice-set such that each item is connected by a fixed number of N symmetric neighbors. This is an instance of the maximum weight b-matching problem (Papadimitriou and Steiglitz 1998) . Exact algorithms exist for solving this problem, such as the blossom algorithm but these are impractical for large-scale applications (Edmonds 1965). Consequently, we use an approximate algorithm based on max-product message passing to find a b-matching (Jebara and Shchogolev 2006). Given a graph on stimuli that is a b-matching, proposals for d-MCMCP can be made in a variety of ways. The selected proposal method is held constant throughout the experiment. The most straightforward proposal method, which we use here, is to choose a proposal uniformly from all *b* neighbors, where the value of *b* is chosen at the experimenter's discretion, leaving a small probability of choosing uniformly from all of the other items (Hsu et al. n.d.).

Deleted: 2011

When the chosen measure of similarity and what's "nearby" matches the notion of nearby in a human respondent's preference space (and presumably profiles with more overlapping attributes will be more similar in preference) then the d-MCMCP algorithm will move away from profiles lower in preference and towards exploring item profiles that are higher in preference. If the definition of similarity is not matched to psychological preference similarity, then the algorithm is still guaranteed to converge but somewhat more slowly. In practice, our simulations below show that even when no neighbors are used, and all proposals are chosen uniformly from all possible profiles, convergence rates are not much slower (see simulations below).

Summary of implementing d-MCMCP. Figure 1 illustrates the sequence of steps involved in implementing d-MCMCP using the example domain of wine profiles, which we will use in our simulations and human experiments. The first step is to create a choice set of product profiles, over which the distribution of relative preference strengths will be estimated (figure

1a). The second step is to compute the similarity between all possible pairs of profiles (figure 1b). The third step is to feed these similarities into a b-matching algorithm to construct a symmetric graph of connections with fixed number of neighbors (figure 1c). The neighboring connections on this graph are then used to define the proposal distribution used in d-MCMCP. Finally, a d-MCMCP experiment is conducted where respondents are asked to choose between two profiles, corresponding to a current and proposal state (figure 1d). (These states are unlabelled for the respondent and are only kept track of in the computer.) The profile chosen by the participant becomes the new current state, and a new proposal is chosen for the next trial. The chain of chosen states will then converge to being samples from the distribution of preferred profiles. In practice, multiple d-MCMCP chains are interwoven randomly among trials. This has the advantages of allowing for multiple chains with multiple random starting states, and also the respondent does not see the same previously chosen state over and over.

Simulations

Here we apply d-MCMCP to simulated choice data to assess its ability to recover the pre-defined distribution of preference strengths over the choice set. For our simulations we will use the domain of wine preferences. We simulate choices based on relative preference strengths that are constructed from pre-defined attribute level utilities. Note, our simulation results are not specific only to wine attributes and would apply as readily to any domain that has similar attribute level structures. As mentioned above, we will compare the convergence results from the d-MCMCP method to that of the naïve alternative of estimating preferences from randomly presented pair-wise choices. By comparing d-MCMCP with the more intuitively transparent pair-wise method, we aim to demonstrate how d-MCMCP works, and

when it may be most useful. In our simulations, we first compare two example sets of attribute utilities, one in which attribute utilities are independent, and the other where there are correlations among attribute utilities. Here, we will examine how d-MCMCP and random pair-wise methods are able to recover the original distribution of profile utilities as a function of the number of trials. Next, for a fixed number of trials, we will apply d-MCMCP and random pair-wise methods to six different preference strength distributions, arising from six respective sets of pre-defined attribute utilities. We show that the advantage of the d-MCMCP method lies in its ability to recover utility distributions that are more selective (preferences are clustered mainly among a small number of items in the choice set).

Two Attribute Sets and Varying Proposals

Stimuli. For the profiles in our choice set, we used 100 real wines that are commonly sold in a major UK supermarket. For our simulation work we chose to characterize these wines in terms of the attributes of price (cheap 0-£5, medium-£5-10, expensive-£10+), color (Red, Rose, White), and Country (Argentina, Australia, Chile, France, Germany, Italy, New Zealand, Portugal, South Africa, Spain, USA). Correlations between attributes in our list of 100 wines did not exceed .25.

Defining Preference Strength Distributions. A preference strength distribution is obtained by first defining utilities on attribute levels. Then based on these attribute utilities, preference strengths can be calculated for each wine profile in the choice set. In particular, for a set of *N* wines, *W1, W2, W3…WN*, each wine can be represented as a binary vector with indices corresponding to all the attribute levels represented in its profile. The binary vector for each wine will contain a 1 for all attribute levels which describe the wine and 0 otherwise. Figure 2a shows an example of such a binary vector for a cheap, French, red wine. The profile utility *U(Wk)* for the *k'th* wine *Wk* is calculated based on the wine's attribute and the defined attribute utilities plus random noise.

Here we examine two different wine distributions of utilities over profiles. Our first distribution of profile utilities is constructed from linear attribute utilities. Linear attribute utilities can be represented by a vector *A1* and we arbitrarily assigned utilities of 1,1,3,2 to the attribute levels of Cheap, French, White, Italian respectively (see figure 2b). Based on these linear attribute utilities, we can calculate the profile utility for wine *W1*, $U_{A1}(W1)$, *as follows:* $U_{A1}(W1) = \sum_i W1_i A1_i + \varepsilon$, where ε is random noise with mean zero. We then obtain the distribution of preference strengths over all wines $U_{A1}(\boldsymbol{W})$, by repeating this for all *100* wines in the choice set $\boldsymbol{W}$*: W1, W2, …W100* (see figure 3a). For illustration purposes preference strengths are shown normalized such that preferences sum to one. In practice, because we are simulating choices based on Luce's choice rule (equation 2), only the relative strengths contribute the probability of choosing a given profile, normalization does not make a difference. (Note, if we were to simulate using a logit choice rule, then the scale would have to be determined).

For our second distribution of profile utilities, we defined correlated attribute utilities, by assigning utilities of 1 and 2, for white-Italian wines, and cheap-French wines respectively. These nonlinear utilities can be represented in a two dimensional matrix *A2* (see figure 2c). Using these attribute utilities we calculate profile utilities over the 100 wines in our choice set as follows: $U_{A2}(\mathrm{W1}) = \sum_j \sum_i W1_i A2_{i,j} W1_j + \varepsilon$, where ε is random noise with mean zero. We use this to obtain the distribution of preference strengths over all wines $U_{A2}(\boldsymbol{W})$ (see figure 3c). Again, the preference utilities are shown normalized to sum to one, though under the assumption of Luce's Choice rule for the binary choice behavior, the normalization does not affect our simulation.

From Preference Strengths to Choices: For a given distribution of preference strengths over all wines, e.g. $U_{A1}(\boldsymbol{W})$, we can then define the probability of choosing one wine, e.g. W1, out of a pair of wines, e.g. W1 and W2, using their respective strengths, e.g. $U_{A1}(W1)$

and $U_{A1}(W2)$, and Luce's Choice rule (equation 2). Now, for any pair of wine profiles, we can simulate a stochastic choice decision. For our noise term, ε, we used Gaussian noise of standard deviation .2 and mean 0.

Simulating d-MCMCP and Random Pair-wise Experiments. For d-MCMCP, we quantified a similarity between each pair of wines in our choice set based on the number of attributes levels the pair of wines had in common. We then input these similarity measures into a b-matching algorithm (Jebara and Shchogolev 2006), which we fixed at 10 neighbors. Proposals were made by randomly choosing uniformly from all profiles with probability $p_{uniform}$, and uniformly from one of the 10 neighbors of the time 1-$p_{uniform}$ of the time. In order to demonstrate the sensitivity of the d-MCMCP algorithm to the method of choosing proposals, we implement our d-MCMCP simulations for values of $p_{uniform}$ set to 0, .33, .66, 1. Note, that d-MCMCP with $p_{uniform}$ =1 is not the same as random pair-wise selection because MCMC chains are still being formed where the previously chosen state becomes the new current state. For the random pair-wise experiment, instead of using current and proposal states, two randomly chosen wines were presented on each trial. For all simulations, estimated distributions of profile utilities (the histogram of choices computed over all wine profiles in our choice set) were evaluated at 20 trial increments from 20-200 trials. Because the simulations are stochastic, we repeated this procedure 100 times. The mean and standard error of these 100 repetitions are shown in the results.

Simulation Results. To quantify how well the methods recovered the actual underlying preference strength distributions, we computed the Pearson's correlation between between the predefined and estimated preference strength distributions over the 100 wines. Figure 3b,d show the mean and standard errors of correlations as a function of number of trials, for d-MCMCP with four different values of $p_{uniform}$ and the random pair-wise method. In general, the value of $p_{uniform}$, does not drastically change performance. The worst performances were

for the extreme values of $p_{uniform}$=0 and 1. Intuitively this makes sense because these represent proposals that are either explore the space not widely enough, or too widely, respectively. However, we note that $p_{uniform}$=1 does not fare much worse than the more optimal $p_{uniform}$=.33 and .66, thus suggesting that the d-MCMCP algorithm may be usefully implemented even without the use of nearest neighbors, and by merely choosing proposals uniformly at random from all profiles in the choice set.

We see that for preferences constructed from linear attribute utilities *A1*, which feature fairly widely distributed preference strengths, d-MCMCP and random pair-wise methods perform roughly similarly in their ability to recover the original preference strength distribution. Note, that d-MCMCP is still guaranteed to eventually converge to the correct distribution of profile utilities (i.e. at sufficient number of trials, the correlation will approach one), whereas a random pair-wise sampling is only guaranteed to converge if the actual distribution were uniform. In contrast to *A1*, for *A2*, where the preference strength distribution is highly peaked and selective for particular wines, d-MCMCP is more effective than random pair-wise choice at recovering the distribution of profile utilities. This makes intuitive sense because d-MCMCP explores the choice-set in a systematic manner, which is what enables d-MCMCP to efficiently explore the range of profiles and focusing on offering choices between profiles with high preference strengths (i.e. profiles that neighbor previously chosen profiles). In contrast, when the distribution approaches that of having no preferences (a uniform distribution of preferences), then a purely random pair-wise sampling method will be equally efficient.

Simulations for Six Sets of Attribute Utilities

Here we repeat the analysis we did above for six different preference strength distributions, chosen for their varying amounts of "selectivity" over wines (see figure 4a). For these simulations, we fix the number of trials at 200 and $p_{uniform}$ =.33. These six distributions

were constructed from the varying combinations of linear and nonlinear attributes shown in Table 1.

---insert Table 1 here---

One way of quantifying the spread of a discrete distribution is it's entropy which is given as:

$$\text{Entropy} = -p(j)\sum_j log(p(j)) \quad (3)$$

Where $p(j)$ is the probability of the value j. Thus, to quantify the spread of a profile utility distribution, we use the normalized profile utilities $U(W)$ for $p(j)$. For this analysis we also use a typically employed method in MCMP experiments, which is to run several chains in the same experiment and average the results of these chains together. This allows us construct chains from multiple starting states. Thus the 200 d-MCMCP trials consisted of five 40-trial chains. The random pair-wise experiment was conducted as before, also for 200 trials. As a way of quantifying the selectivity of the distribution, we calculated the entropy for our six different preference strength distributions. The lower the entropy, the more selective the distribution is. As before, performance was quantified by the correlations between estimated and actual distribution of profile utilities. As with the first set of simulations, we repeated simulations 100 times for each set of attribute utilities. Figure 4b shows the mean and standard error correlation, averaged over 100 repeated simulations, plotted against the entropy of the profile utility distribution (lower entropy means increasing selectivity). Again, our simulation results show that as the profile utility distribution becomes more selective (lower entropy), the advantage of d-MCMCP over random pair-wise samples increases.

Human Experiment

Here we assess the ability of d-MCMCP to estimate real human preference strength distributions using an online experiment. As with the simulations, we compare the results between the two. In order to test how well preference strength distributions are being estimated, we introduce 10 novel wines, which we ask participants to rank order in terms of preference at the end of the experiment. A nonlinear model is fitted to the preference strength distributions obtained from each experiment, and these models are then used to predict the participants' rank ordering of preferences for the 10 novel wines from the two experiments. Spearman's rank correlation between estimated rankings and actual rankings is used as our approximate of how well we were able to estimate a person's wine preference strength distribution. Because our simulations showed that d-MCMCP becomes more advantageous as preferences become more selective, we hypothesize that d-MCMCP will perform better for participants who claim to be expert wine drinkers (i.e. more selective preferences) over those who claim to be non-experts (i.e. less selective preferences).

Method

Participants. 400 online participants were recruited via the Maximiles online panel (http://www.maximiles.co.uk). Participants were randomly assigned to complete an experiment conducted using either the d-MCMCP or the random pair-wise method (200 for each method). Because the experiment was estimated to take at least 20 minutes, participants who took less than this time were discarded from the analysis. This left 187 participants in the d-MCMCP method, and 189 in the random pair-wise method.

Procedure.

Participants were asked to remember the last time they went to a supermarket to shop for wine. To refresh their memory they had to state the type of wine they purchased, the wine's

attributes that were important for them at that time, as well as the occasion they bought the wine for. They were then instructed to imagine repeating that same wine shopping mission at a large national wine chain called 'Jones's Wine Store' and to indicate their preferences in a series of pair-wise choices between wines. We also informed them that the wines shown in this task were both common brands and the store's own-brand wines, which were presented as 'Jones's Own'.

For this human experiment we again use the same 100 commonly purchased wines from a major UK supermarket used in the above simulations. Because we were interested consumer preferences of these wines themselves, in the human experiment we also included information about the Grape, Producer, Wine Maker, Closure type, and Alcohol content as well as the Price, Color, and Country in the wine profiles. We also provided exact prices and the name of the wine. All own-branded wines were labelled as 'Jones's Own' in accordance with the cover story. An example screen shot of the experiment is shown in figure 5.

For the d-MCMCP experiment, we used 10 neighbors and a probability of choosing uniformly from all wines of .33. Also, as done for the 2nd set of simulations, and is the typical protocol in previous d-MCMCP studies (ref), we interweaved five d-MCMCP chains of 40 trials each for a total of 200 trials. This is effectively like running five separate d-MCMCP experiments of 40 trials, but with the questions for these chains interweaved randomly among each other. Along with the advantage of having multiple starting states, this method minimizes the number of times the human participant will see his previously chosen state in the next trial, which would always happen if only one d-MCMCP chain was being constructed. For the random pair-wise method, 200 binary choices between randomly selected pairs were used. Under both methods, the experimental screen for participants looked exactly the same, except the only difference was that participants in the d-MCMCP were answering questions that formed interleaved d-MCMCP chains, which meant they

might see more immediate repetition of the wines they chose in previous trials. An additional change we introduced in the human experiment was that in order to encourage more probabilistic, graded responses, participants were asked to indicate their degree of preference using a slider. Participants were asked "Which wine would you be more likely to purchase?" and asked to position their answer between two ends of a slider that said "Definitely wine A" and "Definitely wine B". The proportion to which the slider was adjusted towards a given wine x′ was taken to be $p_{choice}(x')$.

At the end of both d-MCMCP and random pair-wise experiments, participants were presented with ten additional novel wines that were arbitrarily made up to span the range of possible wine attributes. For this question, five red and five white wines were generated by randomly picking sensible attribute levels from the 100 wines used in the main experiment. When deciding which attribute levels were sensible it was taken care to avoid obvious conflicting combinations of attributes, such as a white grape varieties for a red wine.

After rating these novel wines participants were asked for their self-reported degree of wine expertise by rating their agreement on a 5-point scale (ranging from 'Strongly Disagree', 'Somewhat Disagree', 'Neutral', 'Somewhat Agree', 'Strongly Agree') to the following three statements: 1) I definitely know my favourite wines, where they come from and what grape variety they are. 2) I do not really know much about wine, such as the different regions or grape varieties. 3) I know a lot about wine. Participants' wine expertise were scored by recoding their answers to the numbers -2,-1,0,1,2 for questions 1 and 3 and to 2,1,0,-1,-2 for question 2. Thus, the higher the score, the more wine expertise the participant was deemed.

Model Fitting and Predictions: After the data was collected, distributions of profile utilities, $U(\boldsymbol{W})$, were estimated by computing the histogram of the choices made by each participant. For each participant, based on their estimated $U(\boldsymbol{W})$, nonlinear model fits were

performed using the nonlinear least squares function *nlinfit* in Matlab for the following equation: $\sum_i x_i A1_i + \sum_j \sum_i x_i A2_{i,j} x_{j.} = U(x)$ where the independent variables *x* were the binary vectors describing each wine profile in the choice set, where *A1* and *A2* are the linear and correlated attribute utilities respectively. While we had presented participants the additional attributes of wine maker, producer, exact price, stopper, and alcohol content, in order to limit the number of free parameters in the model fitting, we kept our model parameters to the original simulation attributes of three levels of price, country, and colour, which have *3,11,* and *3* levels respectively. Allowing for both linear and correlated utilities, this leaves us *3+11+3 + 3*11+3*3+11*3* =92 parameters to fit. The fitted parameters for *A1* and *A2* were then used to predict the rank order of each participant's preferences for the 10 novel wines. Spearman's rank correlation was used to quantify the correspondence between the estimated ranking by the fitted model and the actual ranking by the participants.

Results

Participants' rank correlation between predicted and actual rankings for the 10 novel wines were analyzed in a two-way ANOVA looking at the factors of wine expertise (experts versus non-expert) and experimental method (MCMCP versus random pair-wise). Participants were labelled as experts if they obtained an average scored >0 in response to the three wine expertise questions. There were 86 experts and 101 non-experiments in the d-MCMCP experiment, and 88 experts and 101 non-experts in the random pair-wise experiment. An unbalanced two-way ANOVA showed no significant main effects of experimental method ($F(1,372)=.64$, MSE=.06,p=*ns*) or expertise ($F(1,372)=.06$, MSE=.006,p=*ns*) on correlation between predicted and actual rankings for the 10 test wines. However, as predicted from our simulations, we did find a significant interaction between the factors of method and expertise ($F(1,372)=5.07$, MSE= .5, $p=.025$). An individual t-test showed that among expert wine drinkers, the d-MCMCP method had significantly higher

correlation in predictions (M= 0.39, SD= .0033) than the random pair-wise method, (M= .29, SD=.0039), t(172)=2.07, p<.05 (See figure 6). Based on the assumption that experts have more selective preferences, this is consistent with our earlier simulation results that showed that d-MCMCP is more advantageous when preferences are more selective.

Discussion and Future Work

In this paper we have presented the first normative tool for direct measurement of consumer preferences. d-MCMCP is mathematically guaranteed to converge to the real distribution of relative utilities over complex products. It yields results that are consistent with random utility theory, and thus suitable for application to economic models, and assessments of policy and welfare. Although here we demonstrate d-MCMCP on a relatively small (100 item) choice set, with pooled respondent data, d-MCMCP is capable of measuring distributions over much larger sets of items, e.g. thousands of items, as has been done in psychological research (Hsu et al. n.d. ; Sanborn et al. 2010). The ability of d-MCMCP to efficiently explore a large space of profiles alleviates the burden of having to carefully choose a few select profiles for the survey, and potentially reduces the length of study required to explore consumer preferences across a large number of alternatives Also, it requires no a priori assumptions about the possible interactions among attributes. This is in contrast to CAs for which these interactions have to be assumed beforehand and specially included as a compound dimension. Because d-MCMCP does not require a priori assumptions, and because its efficient algorithm can accommodate more factorial combinations of attributes in the choice set, d-MCMCP can capture potential attribute interactions without hypothesizing that they exist beforehand.

Deleted: a

Deleted: recently-developed method in mathematical psychology that can be a useful

Deleted: new

Deleted: measuring

Deleted: of profile utilities over a much large choice set

Deleted: 2011

Deleted: with conjoint methods such as FPCAs

Deleted: often

Deleted: hypothesized

Deleted: In

Deleted: contrast

Deleted: can be accommodated

Deleted: under

Deleted: ,

Deleted: can be captured

The model fitting we presented here, which we applied directly to the estimated distribution of profile utilities, is a slight departure from other choice-based experimental research, which usually fit models to people's choices. Because the primary usage of d-MCMCP is to capture the distribution of profile utilities itself, we chose, here, to fit our models to the estimated preferences directly. However, it is probable that the model fitting can also applied fruitfully to the d-MCMCP choice data. Because d-MCMCP focuses on presenting profile pairs that are most preferable and thus most relevant for determining preferences, it may serve as a natural adaptive algorithm for eliciting choices from the profiles most relevant for assessing preferences and may also be suitable for forms of discrete choice modelling. We leave for future work to see whether there are benefits to fitting the choices instead of preferences directly. It also remains to be explored in future work the many domains for which d-MCMCP may be valuable. Based on our simulation and experimental results, we suggest that the domains in which d-MCMCP may offer particular advantages will be ones where preferences are expected to be selective, with potential attribute weight interactions, and for which one would like to directly obtain a correctly scaled estimate of the relative utility distribution over a set of product profiles. In consumer research, these domains are likely to be those with more complex products with many attributes, e.g. mobile phones or laptops, or those domains where a sizable population of consumers are highly engaged and have complex preferences, e.g. wine or music. Furthermore, d-MCMCP offers a novel method for characterising and classifying consumers into segments with similar preference structures, rather than grouping consumers using self-rated attitudinal or engagement statements such as the expertise ratings used in our study. Segmenting consumers into relatively homogenous groups using their utility profiles over a set of products could help marketers identify unfulfilled needs or estimate the likely demand for new products.

Deleted: established

Deleted: may be a more suitable method compared with other conjoint analyses.

Deleted: may also offer

Reference List

Agarwal, M. K. and P. E. Green (1991), "Adaptive Conjoint Analysis Versus Self Explicated Models: Some Empirical Results," *International Journal of Research in Marketing*, 8 (June), 141-6.

Barker, A. A. (1965), "Monte Carlo calculations of the radial distribution functions for a proton-electron plasma," *Australian Journal of Physics*, 18 (November), 119-33.

Bradley, R. A. (1954), "Incomplete block rank analysis: On the appropriateness of the model of a method of paired comparisons," *Biometrics*, 10 (June), 375-90.

Bradlow, Y. T., Y. Hu, and T. H. Ho (2004), "A Learning-Based Model for Imputing Missing Levels in Partial Conjoint Profiles," *Journal of Marketing Research*, 41 (November), 369-81.

Carmone, F. J. (1987), "ACA system for Adaptive Conjoint Analysis," *Journal of Marketing Research*, 24 (August), 325-7.

Cattin, P. and D. R. Wittink (1989), "Commercial use of conjoint analysis: an update.," *Journal of Marketing*, 53 (July), 91-6.

Clarke, F. R. (1957), "Constant-ratio rule for confusion matrices in speech communication," *Acoustical Society of America*, 29 (February), 715-20.

Edmonds, J. (1965), "Paths, trees and flowers.," *Canadian Journal of Mathematics*, 17 (February), 449-67.

Elrod, T., J. J. Louviere, and K. S. Davey (1992), "An Empirical Comparison of Ratings-Based and Choice-Based Conjoint Models," *Journal of Marketing Research*, 29 (August), 368-77.

Foxall, G. R. (2007), *Explaining Consumer Choice*. London and New York: Palgrave Macmillan.

Foxall, G. R. (2010), *Interpreting Consumer Choice: The Behavioural Perspective Model*. New York: Routledge.

Green, P. E. (1984), "Hybrid Model for Conjoint Analysis: An Expository Review," *Journal of Marketing Research*, 21 155-69.

Green, P. E. (1974), "On the Design of Choice Experiments Involving Multifactor Alternatives," *Journal of Consumer Research*, 1 (September), 61-8.

Green, P. E. and A. M. Krieger (2002), "What's Right with Conjoint Analysis?," *Marketing Research*, 14 (January), 24-7.

Green, P. E. and V. R. Rao (1971), "Conjoint measurement for quantifying judgmental data," *Journal of Marketing Research*, 8 (March), 355-63.

Green, P. E. and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research*, 5 (September), 103-23.

Green, P. E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (October), 3-19.

Hastings, W. K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57 (April), 97-109.

Hauser, J. R. and V. R. Rao (2003), "Conjoint Analysis, Related Modeling, and Applications","," in *Market Research and Modeling: Progress and Prospects: A Tribute to Paul E. Green*, Y. Wind and P. E. Green, eds. Klewer Academic Publishers.

Hopkins, J. W. (1954), "Incomplete block rank analysis: Some taste test results," *Biometrics*, 10 391-9.

Hsu, A., J. B. Martin, T. Griffiths, and A. N. Sanborn (manuscript in preparation), "Extending Markov Chain Monte Carlo with People to categories defined on sets of discrete items,".

Deleted: 2011

Jebara, T. and V. Shchogolev (2006), "B-matching for spectral clustering," *European Conference on MachineLearning, ECML*.

Johnson, R. M. (1987), "Adaptive Conjoint Analysis," *Sawtooth Software Conference on Perceptual Mapping.Conjoint Analysis, and Computer Interviewing*, (Ketchum, ID: Sawtooth Software, Inc), 253-365.

Johnson, R. M. (1974), "Trade-Off Analysis of Consumer Values," *Journal of Marketing*, 11 (May), 121-217.

Kanninen, Bl J. (2002), "Optimal design for multinomial choice experiments," *Journal of Marketing Research*, 39 (May), 214-27.

Louviere, J. J. (1988), *Analyzing decision making: Metric conjoint analysis*. Newbury Park: Sage Publications Inc.

Louviere, J. J., T. N. Flynn, and R. Carson (2010), "Discrete Choice Experiments Are Not Conjoint Analysis," *Journal of Choice Modelling*, 3 (October), 57-72.

Louviere, J. J., D. Hensher, and J. D. Swait (2000), *Stated Choice Methods: Analysis and Application*. Cambridge: Cambridge University Press.

Louviere, J. J., T. Islam, N. Wasi, D. Street, and L. Burgess (2008), "Designing Discrete Choice Experiments: Do Optimal Designs Come at a Price?," *Journal of Consumer Research*, 35 (February), 360-75.

Louviere, J. J., D. Pihlens, and R. Carson (2011), "Design of Discrete Choice Experiments: A Discussion of Issues That Matter in Future Applied Research ," *Journal of Choice Modelling*, 4 (June), 1-8.

Louviere, J. J. and G. Woodworth (1983), "Design and analysis of simulated simulated choice and allocation experiments," *Journal of Marketing Research*, 20 (November), 350-67.

Luce, R. D. (1963), "Detection and recognition.,", R. D. Luce, R. R. Bush, and E. Galanter, eds. New York and London: John Wiley and Sons Inc., 103-90.

Martin, J. B., T. Griffiths, and A. N. Sanborn (2011), "Testing the efficiency of Markov chain Monte Carlo with people using facial affect categories,".

McFadden, D. (1971), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zaremka, ed., 105-42.

Orme, B. K. (2009), *Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research*. Research Pub Llc.

Papadimitriou, C. and K. Steiglitz (1998), *Combinatorial Optimization: Algorithms and Complexity*. New York: Dover.

Park, Y. H., M. Ding, and V. R. Rao (2008), "Eliciting preference for complex products: a web-based upgrading method," *Journal of Marketing Research*, 45 (October), 574.

Pekelman, D. and S. K. Sen (1979), "Measurement and Estimation of Conjoint Utility Functions," *Journal of Consumer Research*, 5 (March), 103-23.

Rao, V. R. (2008), "Developments in conjoint analysis," in *Handbook of Marketing Decision Models*, B. Wierenga, ed. Springer, 23-5.

Rossi, P., G. Allenby, and R. McCulloch (2005), *Bayesian statistic and marketing*. Wiley-Blackwell.

Sanborn, A. N., T. Griffiths, and R. Shiffrin (2010), "Uncovering mental representations with Markov chain Monte Carlo," *Cognitive Psychology*, 60 (March), 63-106.

Sándor, Z. and M. Wedel (2001), "Designing conjoint experiments using manager's prior beliefs," *Journal of Marketing Research*, 38 (November), 430-44.

Sándor, Z. and M. Wedel (2005), "Heterogeneous Conjoint Choice Designs," *Journal of Marketing Research*, 42 (May), 210-8.

Srinivasan, V. (1988), "A Conjunctive-Compensatory Approach to the Self-Explication of Multiattributed Preferences," *Decision Sciences*, 19 (Spring), 295-305.

Toubia, O., J. R. Hauser, and R. Garcia (2007), "Probabilistic polyhedral methods for adaptive choicebased conjoint analysis: Theory and application," *Marketing Science*, 26 (September), 596-610.

Train, K. E. (2003), *Discrete Choice Methods with Simulation*. Cambridge.

Yu, J., P. Goos, and M. Vanderbroek (2009), "Efficient Conjoint Choice Designs in the Presence of Respondent Heterogeneity," *Marketing Science*, 28 (January/February), 122-35.

Table 1: Attribute levels used in simulation of six preference strength distributions

Profile Utility Distribution	Linear attribute levels	Respective utilities	Nonlinear attribute levels	Respective utilities	Entropy
U1	Red, Cheap	1,1	none	none	4.2
U2	Red	1	none	none	3.9
U3	Red	1	French & Cheap, Red & Cheap	1	3.8
U4	none	none	French & Cheap, Australian & Cheap	1,2	2.3
U5	none	none	French & Cheap, White & Italian	2,1	2.2
U6	none	none	French & Cheap	1	1.6

FIGURE 1

STEPS FOR MEASURING PREFERENCES USING MCMCP

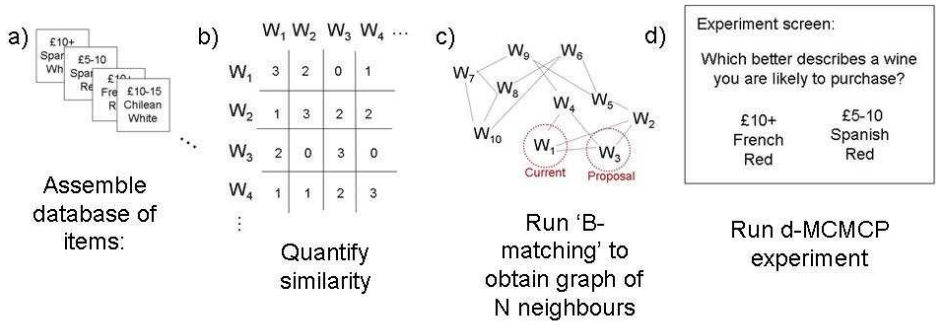

Figure 1: a) Assemble the profiles in the choice set. b) Quantify similarity (using a suitably chosen metric such as number of attributes in common) for all item pairs in the choice set. c) Enter similarity matrix into a B-matching algorithm to obtain a graph where each item is connected by its N neighbors. d) Run d-MCMCP experiment using the neighbors as "nearby" proposals.

FIGURE 2

CALCULATION OF PROFILE UTILTIES FROM MARGINAL ATTRIBUTE UTILITIES

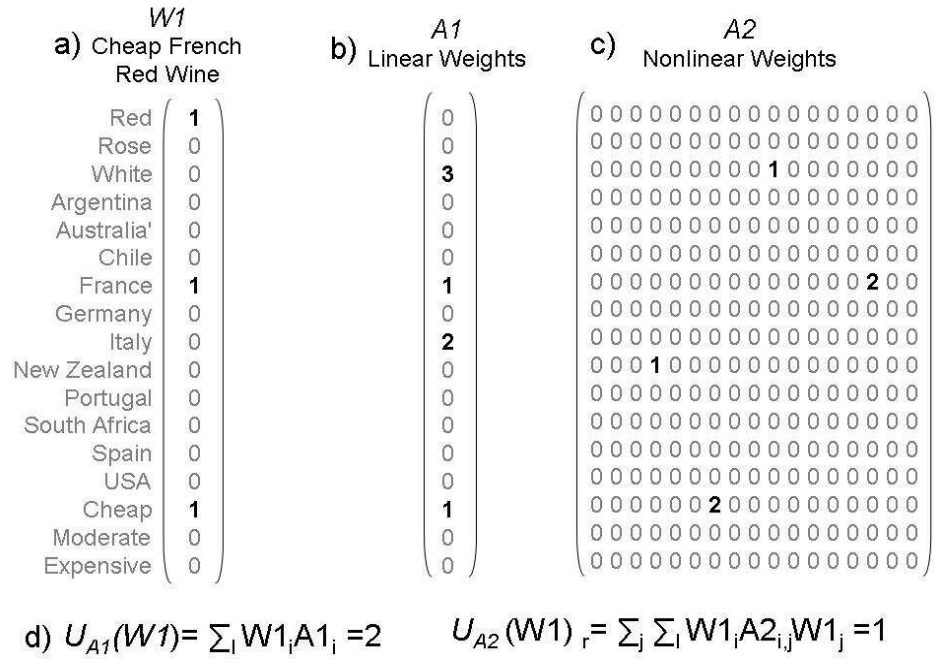

Figure 2: Illustration of how profile utilities are constructed from attribute utilities for our simulations. a) Each wine is represented as a binary vector based on its attribute levels. b) Linear marginal attribute utilities are represented with a vector with weight values assigned to the relevant attribute levels. c) Correlated marginal attribute utilities are represented as a matrix with weight values assigned to the row and column that correspond to the correlated attribute levels. d) Calculation of profile utility for wine shown in a) for under linear and correlated attribute weights, respectively.

FIGURE 3

SIMULATION RESULTS FOR TWO WINE UTILITY PROFILES

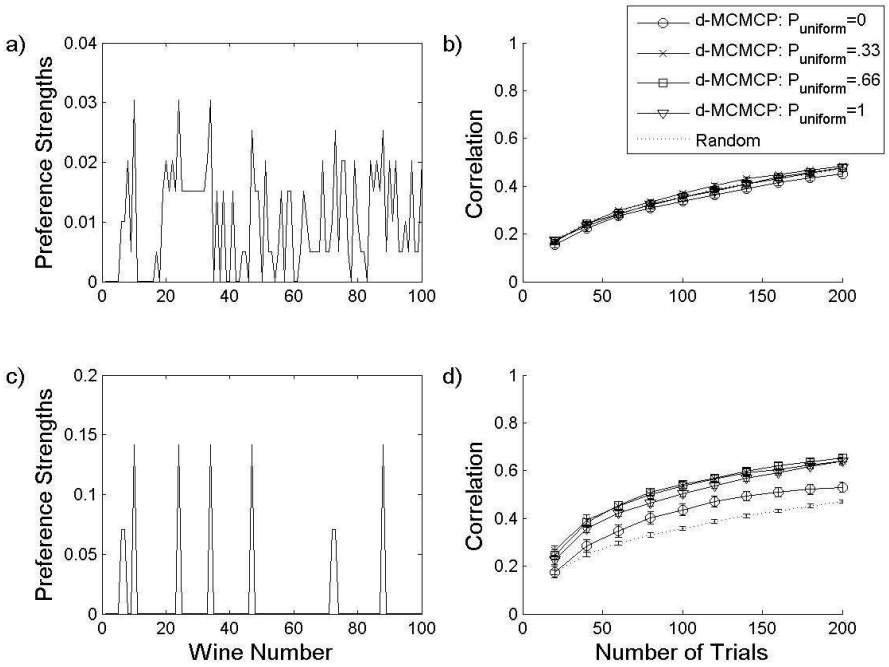

Figure 3: d-MCMCP versus random pair-wise results on two simulated profile utility distributions. a) U1, distribution of profile utilities over 100 wines for linear marginal attribute utilities, normalized to sum to one. b) Correlations between the predefined and estimated preference strength distribution of the 100 wines, as a function of number of choice trials. Different lines correspond to different methods of estimating the profile utility distributions: d-MCMCP implemented with varying probabilities of $p_{uniform}$, the probability of choosing proposals randomly from all possible wines (versus from one of the neighbors), and the random pair-wise method. Each point shows the mean standard error of correlations over 100 stochastic simulation runs. c) U2, distribution of profile utilities over 100 wines for correlated marginal attribute utilities, normalized to sum to one. d) same as b) but for the simulated profile utility distribution shown in c).

FIGURE 4

SIMULATION RESULTS FOR SIX WINE DISTRIBUTION PROFILES

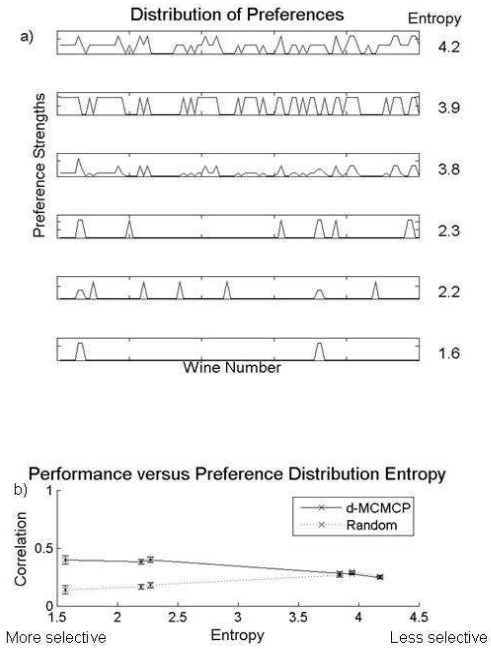

Figure 4: d-MCMCP versus random pair-wise estimates of profile utilities for six different profile utility distributions of varying selectivity. a) Six simulated profile utility distributions over 100 wines. The marginal attribute utilities they were constructed from are listed in table 1. b) Correlation between estimated and actual profile utility distributions plotted as function of entropy of the actual profile utility distribution (higher entropy means more selective distribution). Each point shows the mean standard error of correlations over 100 stochastic simulation runs fixed at 200 trials. d-MCMCP estimates used five chains of 40 trials each with $p_{uniform}$=.33.

FIGURE 5

EXPERIMENT SCREEN SHOT

Figure 5: Example screen shot seen by participants in d-MCMCP and random pair-wise experiments.

FIGURE 6

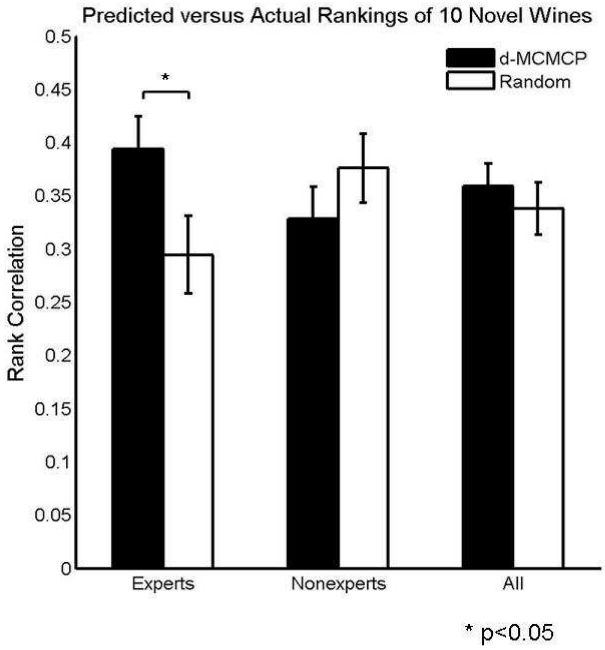

Figure 6: The profile utility distributions estimated from d-MCMCP and random pair-wise methods were fit to a non-linear least squares model. This model fit was then used to make predictions about the participants' rankings of 10 test wines. Participants were split into expert or non-expert groups based on self reported wine expertise. The bar graphs show the mean and standard error of rank correlation between the predicted rankings of 10 novel test wines from both experimental methods, for expert, non-expert and all wine drinkers. For expert wine drinkers, the d-MCMCP method was significantly more predictive of participants' rankings of the 10 test wines compared with the random pair-wise method.