Département de sciences économiques

2004-05

Bayesian Analysis for a Theory of Random Consumer Demand: The Case of Indivisible Goods

McCAUSLAND, William

Université de Montréal

Département de sciences économiques
Université de Montréal
Faculté des arts et des sciences
C.P. 6128, succursale Centre-Ville
Montréal (Québec) H3C 3J7
Canada
http://www.sceco.umontreal.ca
SCECO-information@UMontreal.CA
Téléphone : (514) 343-6539
Télécopieur : (514) 343-7221

Ce cahier a également été publié par le Centre interuniversitaire de recherche en économie quantitative (CIREQ) sous le numéro 10-2004.

This working paper was also published by the Center for Interuniversity Research in Quantitative Economics (CIREQ), under number 10-2004.

ISSN 0709-9231

Bayesian Analysis for a Theory of Random Consumer Demand: The Case of Indivisible Goods

William J. McCausland*

May 31, 2004

Abstract

McCausland (2004a) describes a new theory of random consumer demand. Theoretically consistent random demand can be represented by a "regular" "L-utility" function on the consumption set X.

The present paper is about Bayesian inference for regular L-utility functions. We express prior and posterior uncertainty in terms of distributions over the infinite-dimensional parameter set of a flexible functional form. We propose a class of proper priors on the parameter set. The priors are flexible, in the sense that they put positive probability in the neighborhood of any L-utility function that is regular on a large subset $\bar{X}$ of X; and regular, in the sense that they assign zero probability to the set of L-utility functions that are irregular on $\bar{X}$.

We propose methods of Bayesian inference for an environment with indivisible goods, leaving the more difficult case of infinitely divisible goods for another paper. We analyse individual choice data from a consumer experiment described in Harbaugh et al. (2001).

Key words: Consumer demand, Bayesian methods, Flexible functional Forms, Shape restrictions

*I appreciate comments on earlier versions of this paper by John Geweke, Narayana Kocherlakota and John Stevens. I thank William Harbaugh for providing experimental data. Computations were undertaken using the Bayesian Analysis, Computation and Communication software described in Geweke (1999) and McCausland (2004b). Additional software for computations reported in this paper is available from the author. I alone am responsible for any errors.

1 Introduction

McCausland (2004a) describes a new theory of random consumer demand. Theoretically consistent random demand can be represented by a regular *L-utility*[1] function on the consumption set X. Regular L-utility functions are those that satisfy certain monotonicity and concavity restrictions.

The purpose of this theory of random consumer demand is application to empirical consumer demand problems. To this end, the theory has several desirable features.

1. The representation facilitates inference. Representation theorems identify theoretically consistent random demand with a regular L-utility function, and *vice versa*, so the econometrician can work with regular L-utility functions directly.

2. The representation is parsimonious: a single function on the consumption set describes not only the response of demand to changes in income and prices, as a utility function does in standard consumer theory, but also the distribution of demand on any given budget.

3. The theory is intrinsically stochastic, and so the econometrician can apply the theory directly without recourse to error terms or random preferences. In usual practice, distributions of errors and preferences are given without theoretical justification.

4. The "fit" of an observed choice is measured by the relative desirability of the choice and its feasible alternatives, rather than by some "distance" of the choice to the theoretically optimal option. Varian (1990), in a paper on goodness-of-fit measures, argues for preferring the former to the latter.

5. Unlike standard consumer theory, the new theory does not rule out violations of the usual axioms of revealed preference. In practice, such violations are often observed. The new theory is more forgiving, without being undisciplined.

[1] The term L-utility is meant to invite intuitive comparison with utility while distinguishing it from the usual representation of binary preferences. The L stands for Luce, whose representation in Luce (1959) is similar to the one described in McCausland (2004a).

1.1 A Theory of Random Consumer Demand

We briefly describe the theory proposed in McCausland (2004a). We are in a consumer demand setting with n goods, and the *consumption set* X is[2] $\mathcal{R}^n_+$. A vector $x \in X$ represents a *bundle* of the n goods, and gives quantities of each of the n goods.

A consumer is faced with various non-empty finite[3] subsets of X called *budgets*, and must choose a single bundle from each budget. There is a set $\mathcal{B}$ of all possible budgets, and $(X, \mathcal{B})$ is called the *budget space*.

The primitive concept is that of a *random demand function* p, which assigns to each budget $B \in \mathcal{B}$ a probability distribution p_B on B. Assumptions include stochastic analogues of the classical assumptions of transitivity, monotonicity and convexity. While the classical assumptions apply to binary preference relations, the assumptions of McCausland (2004a) apply to binary choice probabilities (i.e. probabilities of choices from doubleton budgets). Luce's (1959) Choice Axiom relates binary choice probabilities and random demand on choice sets with more than two elements. We can represent theoretically consistent random demand functions by L-utility functions in the set U, defined below, of *regular* L-utility functions.

Definition 1.1 *A function $u : X \to \mathcal{R}$ is* regular *if*

1. *u is non-decreasing, and*
2. *for all prices $w \in \mathcal{R}^n_{++}$ and all incomes $m \in \mathcal{R}^n_+$, u is concave on $\hat{B}(w, m)$, the* classical budget frontier $\{x \in X : w \cdot x = m\}$.

We denote by U the set of regular functions $u : X \to \mathcal{R}$.

A theorem states that for every random demand function p satisfying the assumptions, there exists a regular L-utility function u, unique up to the addition of a constant, such that for every budget B and every element $x \in B$, the probability $p_B(\{x\})$ of choosing bundle x from budget B is given by

$$p_B(\{x\}) = e^{u(x)} \bigg/ \sum_{y \in \hat{B}} e^{u(y)} \;, \tag{1}$$

[2] $\mathcal{R}_+$ is the set of non-negative real numbers

[3] We may consider finite lattices of points in classical budgets of whatever density we like, so this is not a serious restriction. Consumers and econometricians only have a finite set of numbers available to express the quantities of goods they demand or observe. Furthermore, the currency used in transactions is not infinitely divisible.

where the *budget frontier* $\hat{B}$ of budget B is defined by

$$\hat{B} \equiv \{x \in B : \text{there is no } y \in B \backslash \{x\} \text{ such that } y \geq x\}.$$

A second theorem states that the representation is *complete*. For any L-utility function $u \in U$, there exists a random demand function p satisfying the assumptions such that for every budget B, and every bundle $x \in B$, the probability $p_B(\{x\})$ of choosing element x from budget B is given by (1).

We can always make the L-utility function unique by insisting that it take a particular value at a particular point. The representation theorems thus establish an identification of any theoretically consistent random demand function with a regular utility function and *vice versa*. An econometrician can work with regular utility functions rather than random demand functions directly.

In the present paper, we describe Bayesian statistical techniques allowing an econometrician to use the theory and consumer demand data to learn about the behavior of real consumers. The econometrician starts with a prior distribution over L-utility functions, conditions on observed consumer demand data, and obtains a posterior distribution, which identifies the plausible L-utility functions in the light of the data and prior beliefs. The econometrician can go on to make probabilistic predictions about consumer demand on out-of-sample budgets. These predictions take full account of two sources of uncertainty: uncertainty about the L-utility function described by its posterior distribution, and the consumer's random nature governed by a L-utility function.

Learning about a multivariate function that is subject to regularity constraints such as monotonicity and concavity is a familiar problem. Theory imposes constraints on production functions, cost functions, utility functions and indirect utility functions.

There is a large literature on the problem, and the usual approach begins with two choices. The first choice is that of a parametric class of functions, representing either the regular function directly, or a derived function such as a demand function. The second choice consists of constraints on the parameter, which defines a restricted parameter set. The literature identifies two important objectives governing these choices, *theoretical consistency* and *flexibility*. To a large extent, they are competing. Theoretical consistency refers to the extent to which the functions indexed by elements of the restricted parameter set are regular over their domain. If they are regular throughout the domain, we have *global theoretical consistency*. If they are regular at a point, we have *local theoretically consistency*. Flexibility refers to the variety of functions indexed by elements of the restricted parameter set, and it too may be more or less global, depending on how large is the subset of the domain where the relevant flexibility properties hold.

The Constant Elasticity of Substitution (CES) class of utility functions, with non-negativity constraints on its parameters, is globally theoretically consistent but not very flexible. Popular "flexible functional form" classes of demand functions include the translog and Almost Ideal Demand System (AIDS) models. These classes are locally flexible in the sense that with appropriate choice of their parameters they can achieve arbitrary elasticities at a given point in their domain. However, they are not globally theoretically consistent.

There are at least two approaches to approximating regular functions as finite sums of basis functions spanning the space of continuous functions. Gallant (1981) uses a Fourier expansion. Barnett and Jonas (1983) and Geweke and Petrella (2000) use a multivariate Müntz-Szasz expansion. In theory, these approaches offer flexibility and regularity on an arbitrarily large compact subset of the consumption set. The practical problem here is one of estimation. The subset of the parameter space for which the utility function is regular is small and irregularly shaped. Even the task of ascertaining the regularity of a particular utility function is difficult.

We also use finite sums of basis functions to approximate regular functions, and our approach is most similar to that of Geweke and Petrella (2000), in that basis functions are monomials of a transformation of the consumption set. While they use a power transformation, we use a log transformation.

Section 2 is about the approximation of regular L-utility functions by elements in a parametric class of L-utility functions. Following Geweke and Petrella (2000), we apply a result from Evard and Jafari (1994) to show that any regular L-utility function can be arbitrarily well approximated on a hyper-rectangle $\bar{X}$ by a L-utility function in our parametric class that is regular on $\bar{X}$.

In Section 3, we discuss prior distributions over our parametric class of L-utility functions. We show that these priors are proper, which is essential for Bayesian model comparison using Bayes factors. We also show that they are flexible, in the sense that they put positive prior probability in the neighborhood of any L-utility function that is regular on $\bar{X}$; and regular, in the sense that they assign zero probability to the set of L-utility functions that are irregular on $\bar{X}$.

In Section 4, we discuss a demand environment with indivisible goods. The case of infinitely divisble goods is more difficult and is addressed in another paper. We obtain data distributions for quantities demanded given observed prices and income and unobserved L-utility parameters.

In Section 5, we present an empirical application of our theory and econometric methods. We analyse individual choice data from a consumer experiment described in Harbaugh et al. (2001).

We conclude in Section 6.

2 A Parametric Class of L-Utility Functions

Our immediate objective is Bayesian learning, from consumer demand data, about plausible L-utility functions. For this, we require a prior distribution over L-utility functions. Together with data distributions describing random consumer demand given budgets and a L-utility function, we obtain a posterior distribution over L-utility functions.

We propose a flexible infinite-dimensional parametric class of L-utility functions. We will express prior and posterior uncertainty about L-utility functions using probability distributions over the parameter set.

Ideally, we would like our prior over L-utility functions to assign positive probability to any $\|\cdot\|_\infty^X$-neighborhood[4] of any regular L-utility function, and zero probability to the set of irregular L-utility functions. This would require a parametric class containing functions in every $\|\cdot\|_\infty^X$-neighborhood of every regular L-utility function. We settle for a parametric class containing functions in every $\|\cdot\|_\infty^{\bar{X}}$-neighborhood of every regular L-utility function, where $\bar{X}$ is a *restricted consumption set*, defined below:

Definition 2.1 *A set $\bar{X} \subseteq X$ is a* restricted consumption set *if there exist positive $\bar{x}_1, \ldots, \bar{x}_n$ such that*

$$\bar{X} = [0, \bar{x}_1] \times \ldots \times [0, \bar{x}_n].$$

We condition on prices and income, so we can choose $\bar{x} \equiv (\bar{x}_1, \ldots, \bar{x}_n)$ as a function of them. In particular, we can choose $\bar{x}_i > \max_t m_t / w_{it}$ for all $i \in \{1, \ldots, n\}$ to ensure that all feasible bundles lie in $\bar{X}$.

For all $x \in \mathcal{R}^n$ and all multi-indices[5] $\iota \in \mathcal{N}^n$, define $x^\iota \equiv \prod_{i=1}^n x_i^{\iota_i}$. Multi-

[4] For any set A, and any function $f : A \to \mathcal{R}$, we define $\|f\|_\infty^A \equiv \sup_{x \in A} |f(x)|$.

[5] $\mathcal{N} \equiv \{0, 1, \ldots, \infty\}$

variate generalizations of the Stone-Weierstrass theorem tell us that the set $\{x^\iota : \iota \in \mathcal{N}^n\}$ of monomials in x spans the space of continuous functions on the compact set $\bar{X}$. Unfortunately, these monomials do not resemble regular L-utility functions: with positive coefficients, they are convex on classical budget frontiers and with negative coefficients, they are non-increasing. We instead consider the set of monomials $\{[\phi(x)]^\iota : \iota \in \mathcal{N}^n\}$, where ϕ is the following transformation.

We suppose $\bar{x}$ is fixed. Let $\xi \in \mathcal{R}^n_{++}$ and $x^* \in \mathcal{R}^n$ such that $0 < x^* < \bar{x}$ and define $\phi : X \to \mathcal{R}^n$ as

$$\phi(x) = \left(\log\left(\frac{x_1 + \xi_1}{x_1^* + \xi_1}\right), \ldots, \log\left(\frac{x_n + \xi_n}{x_n^* + \xi_n}\right)\right) \quad \forall x \in X. \tag{2}$$

Vectors ξ and x^* are fixed constants that the econometrician choses in advance for computational convenience. They may depend on prices and income, which give useful information about quantity scales, but they do not vary from one L-utility function to another.

The set of monomials in $\phi(x)$ has some useful properties.

1. The positivity of ξ ensures that $\phi(\bar{X})$ is compact, ϕ is invertible on $\bar{X}$, and ϕ^{-1} is uniformly bounded and continuous on $\phi(\bar{X})$. All are required for our approximation result.

2. It includes, for all $i \in \{1, \ldots, n\}$, the regular L-utility function $\log((x_i + \xi_i)/(x_i^* + \xi_i))$.

3. For small values of ξ_i, linear combinations $\sum_{i=1}^n \lambda_i \log((x_i+\xi_i)/(x_i^*+\xi_i))$ of these regular functions approximate Cobb-Douglas L-utility functions of arbitrary "expenditure shares" $\lambda_i / \sum_{j=1}^n \lambda_j$. Higher order monomials in $\phi(x)$ allow changes in "expenditure shares" with prices and income.

4. The constant x^* establishes a reference value around which the monomials of first and second order are nearly "orthogonal".

We approximate L-utility functions as polynomials in $\phi(x)$. We will fix the order $\{\iota^{(k)}\}_{k=1}^\infty$ of multi-indices $\iota^{(k)} \in \mathcal{N}^n$, establishing an order for the monomials in $\phi(x)$. The parameter set is the union $L = \bigcup_{K=1}^\infty \mathcal{R}^K$ of different dimensional subsets. An element $\lambda \equiv (\lambda_1, \ldots, \lambda_K) \in L$ gives the coefficients of the first K monomials.

Definition 2.2 *Define the function* $u : X \times L \to \mathcal{R}$ *by*

$$u(x;\lambda) \equiv \sum_{k=1}^{K} \lambda_k [\phi(x)]^{\iota^{(k)}} = \sum_{k=1}^{K} \lambda_k \prod_{i=1}^{n} \left[\log\left(\frac{x_i + \xi_i}{x_i^* + \xi_i} \right) \right]^{\iota_i^{(k)}} \qquad \forall x \in X,\ \forall \lambda \in L,$$

where K *is the length of* λ.

The unknown parameters are the number of terms K and the coefficients $(\lambda_1, \ldots, \lambda_K)$.

We now identify some important subsets of the parameter set L. One is the subset whose elements index the regular L-utility functions.

Definition 2.3 *For every* $K \in \{1, \ldots, \infty\}$, *define* $\Lambda^K \equiv \{\lambda \in \mathcal{R}^K : u(\cdot, \lambda) \in U\}$, *and* $\Lambda \equiv \bigcup_{K=1}^{\infty} \Lambda^K$.

For every restricted consumption set $\bar{X}$, we define the set $U_{\bar{X}}$ as the set of L-utility functions that are regular[6] on $\bar{X}$ and $\Lambda_{\bar{X}}$ as the subset of parameters indexing L-utility functions that are regular on $\bar{X}$.

Definition 2.4 *For every* K *and every restricted consumption set* $\bar{X}$, *define* $\Lambda^K_{\bar{X}} \equiv \{\lambda \in \mathcal{R}^K \colon u(\cdot, \lambda) \in U_{\bar{X}}\}$ *and* $\Lambda_{\bar{X}} \equiv \bigcup_{K=1}^{\infty} \Lambda^K_{\bar{X}}$.

Note that for every restricted consumption set $\bar{X}$ and every $K \in \{1, \ldots, \infty\}$, $U \subset U_{\bar{X}}$, $\Lambda^K \subset \Lambda^K_{\bar{X}} \subset \mathcal{R}^K$ and $\Lambda \subset \Lambda_{\bar{X}} \subset L$.

2.1 Results

We now present two intermediate results used in Section 3. The following result is also interesting in itself, since the convexity of the parameter subsets $\Lambda^K_{\bar{X}}$ is convenient for posterior simulation.

Result 2.1 *For every* $K \in \{1, \ldots, \infty\}$ *and every restricted consumption set* $\bar{X}$, Λ^K *and* $\Lambda^K_{\bar{X}}$ *are convex cones.*

[6] For every restricted consumption set $\bar{X}$, the function $u \colon X \to \mathcal{R}$ is *regular on* $\bar{X}$ if it is non-decreasing on $\bar{X}$, and for all prices $w \in \mathcal{R}^n_{++}$ and all incomes $m \in \mathcal{R}^n_{+}$, it is concave on $\hat{B}(w,m) \cap \bar{X}$. Recall the definition of the classical budget frontier $\hat{B}(w,m)$ in Definition 1.1.

Proof. U and $U_{\bar{X}}$ are closed under addition and positive scalar multiplication. Therefore Λ^K and $\Lambda^K_{\bar{X}}$ are convex cones. □

The following approximation result tells us that for any restricted consumption set $\bar{X}$, we can approximate any twice continuously differentiable function arbitrarily closely on $\bar{X}$ by a L-utility function in our parametric class. Significantly, we can take the approximating L-utility function to be regular on $\bar{X}$.

The result is similar to a result in Geweke and Petrella (2000). Differences arise because their notion of regularity and their transformation (analogous to ϕ, defined by equation 2) are different. Those authors recognized the significance of a result by Evard and Jafari (1994) on the simultaneous approximation of a function and its derivatives for guaranteeing the regularity of the approximating function.

Result 2.2 (Approximation) *For every restricted consumption set $\bar{X}$, every twice continuously differentiable $u \in U$, and every $\epsilon > 0$, there exists a $\lambda \in \Lambda_{\bar{X}}$ such that*

$$\|u(\cdot;\lambda) - u(\cdot)\|_\infty^{\bar{X}} < \epsilon. \tag{3}$$

Proof. Choose restricted consumption set $\bar{X}$, twice continuously differentiable L-utility function $u \in U$, and $\epsilon > 0$.

We first define a function $\hat{u} : \bar{X} \to \mathcal{R}$, close to u and more convenient to approximate, due to its strict monotonicity and its strict concavity on classical budget frontiers. We will approximate $\hat{u}$, and show that the approximation of $\hat{u}$ is sufficiently close to u.

$$\hat{u}(x) \equiv u(x) + \frac{\epsilon}{2} \prod_{i=1}^{n} \left(\frac{x_i}{\bar{x}_i} \right)^{1/2} \quad \forall x \in \bar{X}.$$

Since u is non-decreasing on $\bar{X}$, concave on all classical budget frontiers, and twice continuously differentiable on $\bar{X}$, $\hat{u}$ is increasing on $\bar{X}$, strictly concave on all classical budget frontiers, and twice continuously differentiable on $\bar{X}$. Also,

$$\|\hat{u}(\cdot) - u(\cdot)\|_\infty^{\bar{X}} = \frac{\epsilon}{2}. \tag{4}$$

A direct corollary of Corollary 3 of Evard and Jafari (1994) is that for every twice continuously differentiable function $f : \bar{X} \to \mathcal{R}$, and every $\epsilon' > 0$, there exists a polynomial $p : \bar{X} \to \mathcal{R}$ such that for all $i, j \in \{1, \ldots, n\}$ and all $x \in \bar{X}$,

$$|f(x) - p(x)| < \epsilon', \quad \left|\frac{\partial f}{\partial x_i} - \frac{\partial p}{\partial x_i}\right| < \epsilon', \text{ and } \quad \left|\frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 p}{\partial x_i \partial x_j}\right| < \epsilon'.$$

The transformation function ϕ (defined in (2)) has an inverse ϕ^{-1} which is twice continuously differentiable on $\phi(\bar{X})$. Therefore $\hat{u} \circ \phi^{-1}$ is also twice continuously differentiable on $\phi(\bar{X})$. Furthermore, $\phi(\bar{X})$ is compact.

The corollary implies that for all $\epsilon' > 0$, there exists a $\lambda \in L$ such that for all $i, j \in \{1, \ldots, n\}$ and all $z \in \phi(\bar{X})$,

$$\left|\sum_{k=1}^{K} \lambda_k z^{\iota^{(k)}} - (\hat{u} \circ \phi^{-1})(z)\right| < \epsilon',$$

$$\left|\frac{\partial}{\partial z_i} \sum_{k=1}^{K} \lambda_k z^{\iota^{(k)}} - \frac{\partial}{\partial z_i}(\hat{u} \circ \phi^{-1})(z)\right| < \epsilon'$$

and

$$\left|\frac{\partial^2}{\partial z_i \partial z_j} \sum_{k=1}^{K} \lambda_k z^{\iota^{(k)}} - \frac{\partial^2}{\partial z_i \partial z_j}(\hat{u} \circ \phi^{-1})(z)\right| < \epsilon',$$

where K is the length of λ.

The function ϕ maps $\bar{X}$ to $\phi(\bar{X})$, and therefore for all $i, j \in \{1, \ldots, n\}$ and all $x \in \bar{X}$,

$$\left|\sum_{k=1}^{K} \lambda_k [\phi(x)]^{\iota^{(k)}} - \hat{u}(x)\right| = |u(x; \lambda) - \hat{u}(x)| < \epsilon',$$

$$\left|\frac{\partial}{\partial x_i} u(x; \lambda) - \frac{\partial}{\partial x_i} \hat{u}(x)\right| < \epsilon' M_1$$

and

$$\left|\frac{\partial^2}{\partial x_i \partial x_j} u(x; \lambda) - \frac{\partial^2}{\partial x_i \partial x_j} \hat{u}(x)\right| < \epsilon' M_2,$$

where M_1 and M_2, derived from uniform bounds on the derivatives of ϕ on $\bar{X}$, do not depend on x.

We can choose ϵ' such that for all $x \in \bar{X}$,

$$|u(x; \lambda) - \hat{u}(x)| < \frac{\epsilon}{2}, \tag{5}$$

$$\frac{\partial u(x;\lambda)}{\partial x} > 0, \tag{6}$$

and

$$v' \frac{\partial^2 u(x;\lambda)}{\partial x \partial x'} v < 0 \quad \forall v \in \mathcal{R}^n \backslash \mathcal{R}^n_+. \tag{7}$$

Equations 4 and 5 and the triangle inequality guarantee that (3) holds. Equations 6 and 7 guarantee that $\lambda \in \Lambda_{\bar{X}}$. $\square$

3 A Class of Priors over L-Utility Functions

We describe here a class of prior densities $f(K,\lambda)$ over the unknown parameters K and λ of the unobserved L-utility function. Each prior $f(K,\lambda)$ has the following desirable features.

- For each K, the conditional prior density $f(\lambda|K)$ is proper and can be evaluated to within a multiplicative normalization constant. This facilitates posterior simulation conditional on K and observed data y, and the approximation of partially marginalized[7] likelihoods $f(y|K)$, for various values of K. We can use the $f(y|K)$ to approximate the posterior probabilities $f(K|y)$ and to do model averaging over K. We can also approximate the fully marginalized likelihood $f(y)$, which we can use to compare the unconditional (on K) model with competing models.

- The prior is *flexible*, in the sense that it assigns positive probability to the $\|\cdot\|^{\bar{X}}_{\infty}$-neighborhood of every regular L-utility function

- The prior is *regular*, in the sense that it assigns zero probability to the set of L-utility functions that are not regular on $\bar{X}$.

Another feature of $f(K,\lambda)$, which is not unambiguously desirable since it is restrictive, is a scale invariance property. In the next section, we express an L-utility function as the product of normalized L-utility function and a multiplicative constant. The scale invariance property is the a priori independance of the normalized function and the multiplicative constant. This independance is computationally convenient and facilitates prior elicitation.

[7]The likelihood has been marginalized with respect to λ, but not K.

We show two results. The first is relevant to the propriety of the priors. The other, building on the approximation result (Result 2.2) of Section 2, establishes flexibility.

3.1 A Class of Priors

Recall that the regular parameter set is $\Lambda_{\bar{X}} = \bigcup_{K=1}^{\infty} \Lambda_{\bar{X}}^K$, and that (Result 2.1) for all K, $\Lambda_{\bar{X}}^K$ is a convex cone. Otherwise $\Lambda_{\bar{X}}$ is a highly irregular set and presents a challenge for prior elicitation.

We consider prior distributions on $\Lambda_{\bar{X}}$ with densities f that can be expressed as

$$f(\lambda) = f_0(u(\bar{x};\lambda)) \sum_{K=1}^{\infty} \pi_K f_K(\lambda|u(\bar{x};\lambda)),$$

where

- $\bar{x} \equiv (\bar{x}_1, \ldots, \bar{x}_n)$ is the far corner of the restricted consumption set $\bar{X} = [0, \bar{x}_1] \times \ldots \times [0, \bar{x}_n]$.
- $f_0(\cdot)$ is any proper density with support $\mathcal{R}_+$. It gives the distribution of $u(\bar{x};\lambda)$, the value of L-utility at $\bar{x}$.
- $\{\pi_K\}_{K=1}^{\infty}$ is sequence of real scalars, with $\sum_{K=\kappa}^{\infty} \pi_K \geq 0$ for all $\kappa \in \{1, \ldots, \infty\}$, and $\sum_{K=1}^{\infty} \pi_K = 1$. For each K, π_K gives the prior probability that the number of basis functions is K.
- For each K such that $\pi_K > 0$, $f_K(\lambda|\bar{u})$ is a density satisfying the following conditions:
  - Its support is the set valued function $\Lambda_{\bar{X}}^K(\bar{u})$ given by

    $$\Lambda_{\bar{X}}^K(\bar{u}) \equiv \{\lambda \in \Lambda_{\bar{X}}^K : u(\bar{x},\lambda) = \bar{u}\} \qquad \forall \bar{u} \in \mathcal{R}_+,$$

    the set of all λ vectors of length K such that $u(\cdot;\lambda)$ is regular on the restricted consumption set $\bar{X}$ and $u(\bar{x};\lambda) = \bar{u}$.
  - The following scale invariance property holds:

    $$f_K(\lambda|\bar{u}) = \bar{u}^{-(K-1)} f_K\left(\frac{\lambda}{\bar{u}}|1\right)$$

The density $f_K(\lambda|1)$ gives (indirectly) the conditional distribution of the normalized L-utility function $u(\cdot;\lambda)/u(\bar{x};\lambda)$ given K.

A key restriction is the scale invariance property, and it implies that the L-utility level $u(\bar{x};\lambda)$ at $\bar{x}$ and the normalized L-utility function $u(\cdot;\lambda)/u(\bar{x};\lambda)$ are independent.

The normalized L-utility $u(\cdot;\lambda)/u(\bar{x};\lambda)$ contains information about the location of modes of choice distributions, and how they respond to changes in prices and income. A natural strategy for eliciting a prior over the normalized L-utility is to put a prior over economically relevant quantities such as the direction of the gradient at various points in the restricted consumption set, thereby inducing an implicit prior over λ. In this way, one can elicit a prior without knowing the details of the L-utility parameterization.

For a given value of the normalized L-utility function, the value $u(\bar{x};\lambda)$ can be interpreted as governing the degree of rationality of the consumer. As it approaches infinity, all choice distributions concentrate on choices that maximise the normalized L-utility. At zero, all choice distributions are uniform over the choices on the budget frontier.

It is easy to ensure that the prior is proper. The first result of the next section is that for every K, the volume of the support $\Lambda^K_{\bar{X}}(1)$ of $f_K(\cdot|1)$ is finite. By choosing a bounded positive function for $f_K(\cdot|1)$, we can ensure that it is finitely integrable on $\Lambda^K_{\bar{X}}$ and thus that $f(\lambda|K)$ is proper.

It is not necessary to know the normalization factor of $f_K(\cdot|\bar{u})$ to be able to do posterior simulation by Markov Chain Monte Carlo (MCMC). Its dependence on $\bar{u}$ has a known form, thanks to the scale independence condition. Its dependence on K is unknown in general, but this is not serious. For various values of K, we can do posterior simulation conditional on K and approximate the partially marginalized likelihood $f(y|K)$ using the methods of Gelfand and Dey (1994) or Meng and Wong (1996), which does not require the normalization constant for $f(\lambda|K)$. From there, we can go on to approximate posterior probabilities of various values of K.

The second result of the next section shows that for suitable choice of the sequence $\{\pi_K\}$, the prior is flexible, in the sense that it puts positive probability in the $\|\cdot\|^{\bar{X}}_\infty$-neighborhood of every regular L-utility function.

3.2 Results

We prove two results. The first result is that for every K the volume[8] $\mathrm{Vol}[\Lambda_{\bar{X}}^{K}(1)]$ is finite. The second result shows that for a suitable choice of $\{\pi_K\}_{K=1}^{\infty}$, each of the priors in our class puts positive probability in the $\|\cdot\|_{\infty}^{\bar{X}}$-neighborhood of every regular L-utility function.

Result 3.1 *For every $K > 0$, $\mathrm{Vol}[\Lambda_{\bar{X}}^{K}(1)]$ is finite.*

Proof. Let K be an arbitrary positive integer. We first establish bounds on the values of L-utility on the subset $[x_1^*, \bar{x}_1] \times \ldots \times [x_n^*, \bar{x}_n]$ of the restricted consumption set.

Claim 3.1 *For every $x \in [x_1^*, \bar{x}_1] \times \ldots \times [x_n^*, \bar{x}_n]$ and every $\lambda \in \Lambda_{\bar{X}}^{K}(1)$,*

$$0 \leq u(x; \lambda) \leq 1$$

Proof. The claim follows directly from the monotonicity of $u(\cdot; \lambda)$, the fact that $u(x^*; \lambda) = 0$, and the fact that $u(\bar{x}; \lambda) = 1$. □

To bound the volume of $\Lambda_{\bar{X}}^{K}(1)$, we first establish the non-singularity of a certain matrix. Once we do this, we can derive an expression for a bound in which the inverse of the matrix appears. First we define some important quantities.

Choose scalars $q_1, \ldots, q_n$ such that for every $i \in \{1, \ldots, n\}$,

1. There exists positive integers m_n and m_d such that $q_i = p_{2i}^{m_n}/p_{2i-1}^{m_d}$, where p_i is the i'th prime number, and

2. $\left[\log\left(\frac{(x_i^*+\bar{x}_i)/2+\xi_i}{x_i^*+\xi_i}\right) \Big/ \log\left(\frac{\bar{x}_i+\xi_i}{x_i^*+\xi_i}\right)\right]^{1/K} \leq q_i \leq 1.$

A simple modification of the proof in Rudin (1976) of the denseness of the rational numbers in the reals shows that we can do this. Note that the inequalities $x_i^* < (x_i^* + \bar{x}_i)/2 < \bar{x}_i$ ensure that we are taking the K'th root of a positive real number strictly less than one.

Now define, for all $k \in \{1, \ldots, K\}$,

$$z_k \equiv \left(q_1^k \log\left(\frac{\bar{x}_1 + \xi_1}{x_i^* + \xi_1}\right), \ldots, q_n^k \log\left(\frac{\bar{x}_n + \xi_n}{x_i^* + \xi_n}\right)\right), \qquad x_k \equiv \phi^{-1}(z_k),$$

[8] We use the notation Vol to denote the volume of a set under Lesbesgue measure.

and

$$C \equiv \begin{bmatrix} z_1^{\iota^{(1)}} & \cdots & z_1^{\iota^{(K)}} \\ \vdots & \ddots & \vdots \\ z_K^{\iota^{(1)}} & \cdots & z_K^{\iota^{(K)}} \end{bmatrix}.$$

Note that for all $\lambda \in \Lambda_{\bar{X}}^K$,

$$C\lambda = [u(x_1;\lambda), \ldots, u(x_K;\lambda)]' = [u(\phi^{-1}(z_1);\lambda), \ldots, u(\phi^{-1}(z_K);\lambda)]',$$

and that for all $k \in \{1, \ldots, K\}$, $x_k = \phi^{-1}(z_k) \in [x_1^*, \bar{x}_1] \times \ldots \times [x_n^*, \bar{x}_n]$. Claim 3.1 gives us $(0, \ldots, 0)' \leq C\lambda \leq (1, \ldots, 1)'$.

We now show that C is non-singular.

Claim 3.2 *C is non-singular.*

Proof. C can be written as

$$\begin{bmatrix} (q^{\iota^{(1)}})^1[\phi(\bar{x})]^{\iota^{(1)}} & \cdots & (q^{\iota^{(K)}})^1[\phi(\bar{x})]^{\iota^{(K)}} \\ \vdots & \ddots & \vdots \\ (q^{\iota^{(1)}})^K[\phi(\bar{x})]^{\iota^{(1)}} & \cdots & (q^{\iota^{(K)}})^K[\phi(\bar{x})]^{\iota^{(K)}} \end{bmatrix}$$

$$= \begin{bmatrix} (q^{\iota^{(1)}})^1 & \cdots & (q^{\iota^{(K)}})^1 \\ \vdots & \ddots & \vdots \\ (q^{\iota^{(1)}})^K & \cdots & (q^{\iota^{(K)}})^K \end{bmatrix} \cdot \operatorname{diag}\left([\phi(\bar{x})]^{\iota^{(1)}}, \ldots, [\phi(\bar{x})]^{\iota^{(K)}}\right)$$

We will show that both these factors are non-singular, which will then imply that C is non-singular. The first factor is a Vandermonde matrix, and to establish its non-singularity, it suffices to show that for all $k, l \in \{1, \ldots, K\}$, $k \neq l \Rightarrow q^{\iota^{(k)}} \neq q^{\iota^{(l)}}$. This follows from the fact that there is a unique representation of any rational number as the ratio of two integers with no common factors, and unique prime factorizations of the two integers. The second factor is a diagonal matrix whose elements are non-zero, and so it is also non-singular. Since the two factors are non-singular, so is C. □

We now construct the matrix $\tilde{C}$ whose inverse is a factor of the bound on $\text{Vol}[\Lambda_{\bar{X}}^K(1)]$, and show that it is non-singular. Let

$$c = \left([\phi(\bar{x})]^{\iota^{(1)}}, \ldots, [\phi(\bar{x})]^{\iota^{(K)}}\right)$$

and note that $c\lambda = \sum_{k=1}^K \lambda_k [\phi(\bar{x})]^{\iota^{(k)}} = u(\bar{x}, \lambda)$. Now construct the matrix $\tilde{C}$ by replacing one row of C with c. We take care to choose a row such that c

is not in the subspace spanned by the remaining rows. Let κ be the index of the row replaced by c. Since C is non-singular, $\tilde{C}$ is also non-singular.

We now show that $\Lambda_{\bar{X}}^K(1)$ is a subset of a set whose larger volume we can bound above.

Claim 3.3

$$\Lambda_{\bar{X}}^K(1) \subseteq \left\{\lambda \in \mathcal{R}^K : (\ldots,0,1,0,\ldots)' \leq \tilde{C}\lambda \leq (1,\ldots,1)'\right\},$$

where the 1 *appears at the* κ*'th place of the vector on the left.*

Proof. Let $\lambda \in \Lambda_{\bar{X}}^K(1)$. Then $(\tilde{C}\lambda)_\kappa = 1$. By the construction of the z_k,

$$x_k = \phi^{-1}(z_k) \in [x_1^*, \bar{x}_1] \times \ldots \times [x_n^*, \bar{x}_n] \qquad \forall k \in \{1,\ldots,K\}.$$

By Claim 3.1, $0 \leq u(x_k;\lambda) \leq 1$ for all $k \in \{1,\ldots,K\}$, and so $0 \leq C\lambda \leq (1,\ldots,1)$. Therefore

$$(\ldots,0,1,0,\ldots)' \leq \tilde{C}\lambda \leq (1,\ldots,1)'. \square$$

Now, by Claim 3.3,

$$\begin{aligned}\text{Vol}\left[\Lambda_{\bar{X}}^K(1)\right] &\leq \text{Vol}\left[\{\lambda \in \mathcal{R}^K : (\ldots,0,1,0,\ldots)' \leq \tilde{C}\lambda \leq (1,\ldots,1)'\right] \\ &= \left|\tilde{C}\right|^{-1}. \square\end{aligned}$$

The second result builds on Result 2.2 on the approximation of L-utility functions to show that the prior is flexible and regular.

Result 3.2 *For all twice continuously differentiable and concave $u \in U$, and all priors in the class described in this section, the prior assigns positive probability to any $\|\cdot\|_\infty^{\bar{X}}$-neighborhood of u, and zero probability to the set of L-utility functions that are not regular on $\bar{X}$.*

Proof. Let $u \in U$ be twice continuously differentiable and concave, and let $\epsilon > 0$. By Result 2.2, we can find a $K \in \{1,\ldots,\infty\}$ and a $\lambda^* \in \Lambda_{\bar{X}}^K$ such that $\|u(x;\lambda^*) - u(x)\|_\infty^{\bar{X}} < \epsilon/2$. The prior, strictly positive on $\Lambda_{\bar{X}}^K$, assigns positive probability to the set

$$\Lambda^* = \left\{\lambda \in \Lambda_{\bar{X}}^K : |\lambda_k^* - \lambda_k| < \frac{\epsilon}{4K}\left(\sup_{x\in\bar{X}}\left|\phi(x)^{\iota(k)}\right|\right)^{-1} \quad \forall k \in \{1,\ldots,K\}\right\}.$$

Table 1: Observed Variables

Variable	Description
w_{it}	price of good i at time t
x_{it}	quantity demanded of good i at time t
m_t	income at time t
$w_t \equiv (w_{1t}, w_{2t}, \ldots, w_{nt})'$	time t price vector
$x_t \equiv (x_{1t}, x_{2t}, \ldots, x_{nt})'$	time t quantity vector
$W_T \equiv \begin{bmatrix} w_1 & w_2 & \ldots & w_T \end{bmatrix}$	price matrix
$X_T \equiv \begin{bmatrix} x_1 & x_2 & \ldots & x_T \end{bmatrix}$	quantity matrix
$M_T \equiv \begin{bmatrix} m_1 & m_2 & \ldots & m_T \end{bmatrix}$	income vector

For all $\lambda \in \Lambda^*$, $\|u(\cdot;\lambda) - u(\cdot;\lambda^*)\|_\infty^{\bar{X}} < \frac{\epsilon}{2}$, and therefore $\|u(\cdot;\lambda) - u(\cdot)\|_\infty^{\bar{X}} < \epsilon$. Since $\sum_{k=K}^{\infty} \pi_k > 0$, the prior assigns positive probability to Λ^*, and therefore a positive probability to the $\|\cdot\|_\infty^{\bar{X}}$-neighborhood of u. The fact that the prior assigns zero probability to the set of functions that are not regular on $\bar{X}$ follows trivially from the fact that the support of the prior is $\Lambda_{\bar{X}}$. $\square$

4 Demand Environment

We have a theory of random consumer choices from arbitrary finite subsets of the consumption set. In the usual consumer demand environment, we do not observe budgets directly, but rather prices and income, which then determine budgets. We describe a consumer demand environments with indivisible goods. We give a mapping from prices and income to finite budgets, data distributions on the budgets for a given L-utility function, and the likelihood function for a given data set.

We observe a consumer making demand decisions at times $t = 1, \ldots, T$, and assume that quantity observations are independent. Table 1 lists the variables we observe, and their notation.

4.1 Indivisible Goods Budgets

The choice environment is useful for the analysis of certain consumer experiments, for example those in Sippel (1997), Mattei (2000), and Harbaugh et al. (2001). Goods are indivisible, and feasible budgets typically have a small number of the indivisible unit of each good. Without loss of generality, we can assign to unity the quantity of a single indivisible unit of each good.

The non-negative income m and positive price vector w are otherwise unrestricted. The mapping B^{IG} from income and prices to budgets is given by

$$B^{\mathrm{IG}}(w,m) \equiv \{x \in X : w \cdot x \leq m \text{ and } x \in \mathcal{N}^n\}$$

Given a L-utility function $u(\cdot, \lambda)$ indexed by some $\lambda \in \Lambda_{\bar{X}}$, and values w and m of prices and income, the probability mass function governing choices on the budget $B^{\mathrm{IG}}(w,m)$ is given by

$$f(x|\lambda, w, m) = \begin{cases} e^{u(x;\lambda)} \Big/ \sum_{y \in \hat{B}^{\mathrm{IG}}(w,m)} e^{u(y;\lambda)} & x \in \hat{B}^{\mathrm{IG}}(w,m) \\ 0 & x \notin \hat{B}^{\mathrm{IG}}(w,m) \end{cases},$$

and the likelihood function for a given data set X_T, W_T and M_T is given by

$$L(\lambda; X_T, W_T, M_T) = \prod_{t=1}^{T} \left[e^{u(x_t;\lambda)} \Bigg/ \sum_{\hat{B}(w_t,m_t)} e^{u(x;\lambda)}\, dx \right].$$

The log-likelihood function for the data set is given by

$$\mathcal{L}(\lambda; X_T, W_T, M_T) = \sum_{t=1}^{T} \left[u(x_t;\lambda) - \log \sum_{\hat{B}(w_t,m_t)} e^{u(x;\lambda)}\, dx \right].$$

When we come to Markov Chain Monte Carlo simulation of the posterior distribution, the following result is reassuring.

Result 4.1 (Log-Concavity of Likelihood Functions) *For every $K \in \{1, \ldots, \infty\}$, every restricted consumption set $\bar{X}$, every allowable data set (X_T, W_T, M_T) for the indivisible good case the log-likelihood $\mathcal{L}$ is concave on $\Lambda^K_{\bar{X}}$.*

Proof. Let $K \in \{1, \ldots, \infty\}$, restricted consumption set $\bar{X}$, and data set (X_T, W_T, M_T) be arbitrary. We will show that for every $t \in \{1, \ldots, T\}$, the t'th term of the log-likelihood $\mathcal{L}$ is concave. Then the log-likelihood $\mathcal{L}$, a sum of these terms, is also concave.

Let $t \in \{1, \ldots, T\}$ be arbitrary. The t'th term of the log-likelihood is

$$u(x_t; \lambda) - \log \sum_{x \in \hat{B}(w_t, m_t)} e^{u(x;\lambda)}\, dx. \tag{8}$$

Since $u(x_t; \lambda)$ is a linear function of λ on $\Lambda^K_{\bar{X}}$, it is concave there. We now show that the log term of (8) is convex. Its gradient is given by[9]

$$\begin{aligned} \frac{\partial}{\partial \lambda} \log \sum_{x_t \in \hat{B}(w_t, m_t)} e^{u(x;\lambda)}\, dx &= \frac{\sum_{x_t \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)} \frac{\partial u(x;\lambda)}{\partial \lambda}\, dx}{\sum_{x_t \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)}\, dx} \\ &= \mathrm{E}\left[\left. \frac{\partial u(x_t;\lambda)}{\partial \lambda} \right| \lambda \right], \end{aligned}$$

and its Hessian is given by

$$\begin{aligned} &\frac{\partial^2}{\partial \lambda \partial \lambda'} \log \sum_{x \in \hat{B}(w_t, m_t)} e^{u(x;\lambda)}\, dx \\ =\; & \frac{\sum_{x_t \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)} \left(\frac{\partial u(x;\lambda)}{\partial \lambda} \frac{\partial u(x;\lambda)}{\partial \lambda'} + \frac{\partial^2 u(x;\lambda)}{\partial \lambda \partial \lambda'} \right) dx}{\sum_{x \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)}\, dx} \\ -\; & \frac{\left(\sum_{x_t \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)} \frac{\partial u(x;\lambda)}{\partial \lambda}\, dx \right) \left(\sum_{x_t \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)} \frac{\partial u(x;\lambda)}{\partial \lambda'}\, dx \right)}{\left[\sum_{x_t \in \hat{B}(w_t,m_t)} e^{u(x;\lambda)}\, dx \right]^2} \\ =\; & \mathrm{E}\left[\left. \frac{\partial u(x_t;\lambda)}{\partial \lambda} \frac{\partial u(x_t;\lambda)}{\partial \lambda'} \right| \lambda \right] - \mathrm{E}\left[\left. \frac{\partial u(x_t;\lambda)}{\partial \lambda} \right| \lambda \right] \cdot \mathrm{E}\left[\left. \frac{\partial u(x_t;\lambda)}{\partial \lambda'} \right| \lambda \right] \\ =\; & \mathrm{Var}\left[\left. \frac{\partial u(x_t;\lambda)}{\partial \lambda} \right| \lambda \right] \end{aligned}$$

We use the fact that the Hessian of u with respect to λ is identically equal to zero. The final right hand side expression is positive semidefinite, and so the log term of (8) is convex. □

[9] $\mathrm{E}[\cdot|\lambda]$ and $\mathrm{Var}[\cdot|\lambda]$ are the expectation and variance operators, respectively, for the conditional distribution of x_t given λ.

5 An Empirical Application

The theory of consumer demand in McCausland (2004a) concerns individual, rather than aggregate choice behavior. In this paper, we focus on indivisible goods and assume choices on different budgets are independent. In addition, posterior simulation is computationally practical only for a small number n of goods.

For these reasons, consumer demand experiments like those described in Harbaugh (2001), Mattei (2000), and Sippel (1997) are ideal as a first application of the theory and econometric techniques. These experiments share the following features.

1. Consumers select bundles from several different budgets, in the knowledge that after all decisions are made, exactly one of the budgets will be selected at random, and the consumer will be given their choice from (only) that budget. We can thus plausibly consider choices as being simultaneous or static, rather than dynamic.

2. Goods are consumed on the spot, shortly after choices have been made.

3. Consumers have the opportunity to go back and change earlier choices, before a budget is selected at random. This mitigates criticism that learning during the experiment is a problem.

4. Choices are reliably recorded. We can be fairly confident that measurement error is not a problem.

5. The number of goods, the nature of the indivisibilities, and prices and income are such that the number of possible choices is computationally tractible.

In this paper, we analyze data from the Harbaugh et al. (2001) "GARP for Kids" experiment, undertaken in a study of the development of rational behavior. Subjects are 31 second grade students, 42 sixth grade students and 55 undergraudates. There are two good, chips and juice, in indivisible packages. There are no prices and income as such: subjects are offered a budget of choices directly, and the budgets do not include off-frontier bundles. Figure 1 illustrates the eleven different budgets.

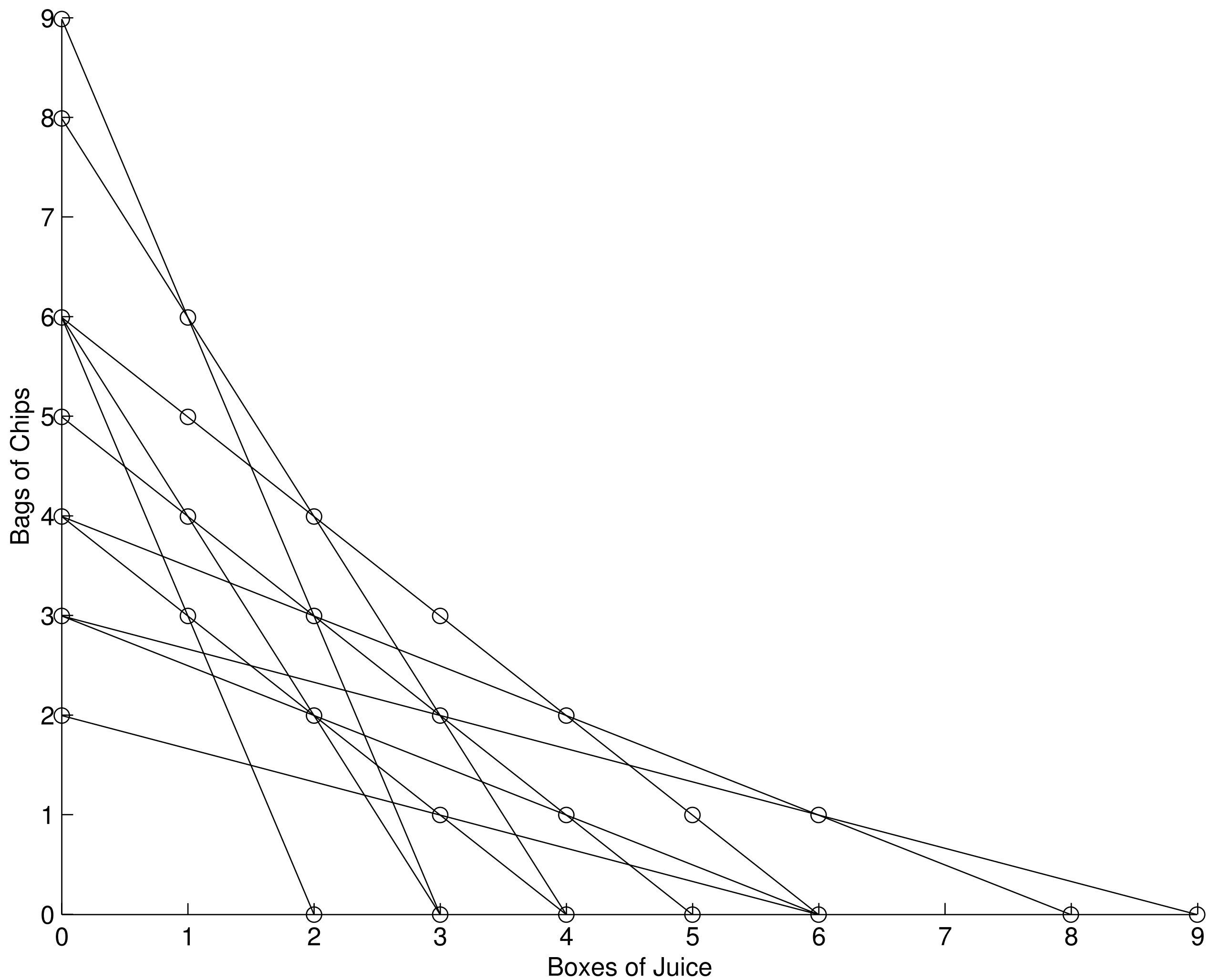

Figure 1: Budgets for the “GARP for Kids” experiment.

5.1 Choice of Constants and Prior

We choose the constants defining the restricted consumption set $\bar{X}$ and the transformation ϕ as $\bar{x} = (10, 10)$, $x^* = (1.0, 1.0)$, and $\xi = (0.1, 0.1)$. Thus $\bar{X} = [0, 10]^2$, which contains all eleven budgets, and the tranformation ϕ is given by

$$\begin{aligned}\phi(x_1, x_2) &= \left(\log\left(\frac{x_1 + \xi_1}{x_1^* + \xi_1}\right), \log\left(\frac{x_2 + \xi_2}{x_2^* + \xi_2}\right)\right) \\ &= \left(\log\left(\frac{x_1 + 0.1}{1.1}\right), \log\left(\frac{x_2 + 0.1}{1.1}\right)\right).\end{aligned}$$

We order the multi-indices $\{\iota^{(k)}\}_{k=1}^{\infty}$ as $(1,0)$, $(0,1)$, $(2,0)$, $(1,1)$, $(0,2)$, $(3,0)$, $(2,1)$, ... and thus the basis functions are ordered ϕ_1, ϕ_2, ϕ_1^2, $\phi_1\phi_2$, ϕ_2^2, ϕ_1^3, $\phi_1^2\phi_2$, ...

The prior on K, the number of basis functions, puts positive prior probability (only) on values of K for which the first K basis functions are those having sums of exponents $\iota_1^{(k)} + \iota_2^{(k)}$ up to some integral order. For $p = 1, 2, \ldots$, we set $\pi_{(p+1)(p+2)/2-1} = 2^{-p}$. Thus $\pi_2 = 1/2$, $\pi_5 = 1/4$, $\pi_9 = 1/8$, etc.

We choose a gamma prior for $u(\bar{x}; \lambda)$, with shape parameter $\alpha = 2$ and inverse scale parameter $\beta = 0.05$. For each K with $\pi_K > 0$, we choose a flat prior for $f_K(\lambda | u(\bar{x}; \lambda)$ on $\Lambda_{\bar{X}}^K(\bar{x}; \lambda))$.

5.2 Simulation

For a given value of K, we simulate finite sequences from two ergodic Metropolis-Hastings Markov chains. The respective stationary distributions are the prior distribution $\lambda | K$ and the posterior distribution $\lambda | K, X_T, W_T, M_T$.

For both Metropolis-Hastings chains, the proposal distribution is a multivariate normal distribution with mean equal to the current value of λ. For prior simulation, the proposal precision (inverse of variance) is a weighted sum of the Hessian of the log prior and a matrix G designed to capture the shape of $\Lambda_{\bar{X}}$. For posterior simulation, the proposal precision is a weighted sum of the same two matrices and the Hessian of the likelihood. The matrix G and the weights are tuned to improve the performance of the chains. Recall that for each K, the set $\Lambda_{\bar{X}}^K$ is convex and that the restriction of the likelihood to it is log-concave. Both are very convenient for posterior simulation.

We approximate partially marginalized likelihoods using the method of Meng and Wong (1996). For several values of K, we draw a sample of

Table 2: Log Marginal Likelihoods for Subjects in "GARP for Kids" Experiment

-7.23	-4.99	-9.42	-4.14	-4.05
-14.56	-9.51	-14.84	-8.13	-15.67
-17.20	-4.15	-17.31	-12.14	-24.99
-20.18	-24.55	-21.18	-4.03	-7.20
-9.42	-4.07	-17.16	-7.18	-8.54
-10.36	-12.92	-4.08	-4.09	-8.98
-11.51	-22.91	-11.76	-12.75	-17.44
-7.16	-11.57	-15.83	-17.19	-4.12
-9.94	-9.83	-22.27	-7.76	-15.15
-19.81	-15.11	-17.62	-10.59	-23.20
-22.69	-4.02	-16.49	-7.21	-20.02

values from the distributions $\lambda|K$ and $\lambda|K, X_T, W_T, M_T$. For each draw of both samples, we evaluate an unnormalized density for the distribution $\lambda|K$ and the likelihood $\mathcal{L}(\lambda; X_T, W_T, M_T)$. The product is an evaluation of an unnormalized density for the distribution $\lambda|K, X_T, W_T, M_T$. The Meng and Wong method gives the ratio of the constants which normalize the two unnormalized densities, which is equal to the value $f(X_T, W_T, M_T|K)$ of the partially marginalized likelihood.

5.3 Results

Table 2 shows results for the "GARP for Kids" experiment. The objective here is not to study the development of rational behavior in children, and so we report results only for the benchmark undergraduate subjects. We report the log marginal likelihood for each subject. Here, the marginal likelihood is the marginal probability that the theory assigns to the sequence of observed choices that a subject makes, for the chosen prior. Standard errors for the numerical approximation of the log marginal likelihoods reported in Table 2 are all less than 0.05. The average log marginal likelihood is -12.48.

To put these quantities in perspective, we consider the average log marginal likelihood arising from various models. The model assigning equal probability to all possible sequences of eleven choices implies a log marginal likelihood of -16.31 for every subject. A model which correctly and with certainty

Table 3: Counts of Numbers of GARP Violations

Number of violations	Experimental subjects	All sequences
0	36	108,846
1	0	0
2	1	140,788
3	5	171,718
4	0	272,978
5	7	438,074
6	1	646,288
7	0	928,790
8	0	1,567,246
9	2	2,081,452
10	1	2,555,030
11	2	3,184,790
Total	55	12,096,000

predicts the behavior of all subjects on all budgets implies a log marginal likelihood of zero for every subject. Any model that assigns probability zero to every sequence featuring at least one violation of the Generalized Axiom of Revealed Preference (GARP) gives a log marginal likelihood of negative infinity to the sequences of the 19 out of 55 subjects who violated the GARP, and therefore an average log marginal likelihood of negative infinity.

We use the data in Table 3 to derive a maximum possible log marginal likelihood of -13.36 for any model assigning equal probabilities to all sequences featuring the same number of violations of the GARP. The second column gives, for the number of GARP violations in the first column, the number of subjects having that number of violations. The third column gives the total number of distinct sequences of eleven choices having that number of violations.

6 Conclusions

We have shown how to do Bayesian Analysis for the theory of random consumer demand in McCausland (2004a). Approximation of utility functions is quite flexible, and so the econometric analysis does not add any serious additional restrictions to random consumer demand that are not theoretically motivated. We provide a class of priors that is proper, practical for the purposes of posterior simulation, and allowing indirect elicitation through priors on economically relevant quantities. We show how to construct a likelihood function for a consumer environment with indivisible goods, deferring the case of infinitely divisible goods to another paper. We have applied the theory of McCausland (2004a) and the econometric techniques of this paper to the analysis of data from a consumer experiment.

References

[1] William A. Barnett, John F. Geweke, and Piyu Yue. Seminonparametric Bayesian estimation of the Asymptotically Ideal Model: The AIM consumer demand system. In William A. Barnett, James Powell, and George E. Tauchen, editors, *Nonparametric and semiparametric methods in econometrics and statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, chapter 6, pages 127–173. Cambridge University Press, Cambridge, New York and Melbourne, 1991.

[2] James O. Berger. *Statistical Decision Theory and Bayesian Analysis, Second Edition*. Springer-Verlag, New York, NY, 1985.

[3] José M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. John Wiley and Sons, Chichester, England, 1994.

[4] Angus Deaton and John Muellbauer. *Economics and Consumer Behavior*. Cambridge University Press, Cambridge, 1980.

[5] W. E. Diewert and T. J. Wales. Flexible functional forms and global curvature conditions. *Econometrica*, 55(1):43–68, 1987.

[6] J. C. Evard and F. Jafari. Direct computation of the simultaneous Stone-Weierstrass approximation of a function and its partial derivatives in banach spaces, and combination with hermite interpolation. *Journal of Approximation Theory*, 78:351–363, 1994.

[7] A. E. Gelfand and D. K. Dey. Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society Series B*, 56:501–514, 1994.

[8] John Geweke. Monte Carlo simulation and numerical integration. In Hans M. Amman, David A. Kendrick, and John Rust, editors, *Handbook of Computational Economics: Volume 1*, chapter 15, pages 731–800. Elsevier Science Publishers, North–Holland, 1996.

[9] John Geweke. Using simulation methods for Bayesian econometric models: Inference, development, and communication. *Econometric Reviews*, 18:1–126, 1999.

[10] John Geweke and Lea Petrella. Inference for regular functions using Müntz approximations. Unpublished Manuscript, 2000.

[11] William Harbaugh, K. Krause, and T. Berry. Garp for kids: On the development of rational choice behavior. *American Economic Review*, 91:1539–1545, 2001.

[12] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[13] Lawrence Lau. Functional forms in econometric model building. In Zvi Gliliches and M. D. Intriligator, editors, *Handbook of Econometrics: Volume 3*, chapter 26, pages 1515–1566. Elsevier Science Publishers, North–Holland, 1986.

[14] Aurelio Mattei. Full-scale real tests of consumer behavior using experimental data. *Journal of Economic Behavior and Organization*, 43:487–497, 2000.

[15] Rosa L. Matzkin. Restrictions of economic theory in nonparametric methods. In R. F. Engle and D. L. McFadden, editors, *Handbook of Econometrics: Volume 4*, chapter 42, pages 2523–2558. Elsevier Science Publishers, North–Holland, 1994.

[16] William J. McCausland. A theory of random consumer demand. Cahiers de recherche du Département de sciences économiques, Université de Montréal, no. 2004-04, 2004.

[17] William J. McCausland. Using the BACC software for Bayesian inference. *Journal of Computational Economics*, 23:201–218, 2004.

[18] X.-L. Meng and W. H. Wong. Simulating ratios of normalizing constants via a simple identity: A theoretical exploration. 6:831–860, 1996.

[19] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys*, 21:1087–1092, 1953.

[20] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993.

[21] Walter Rudin. *Principles of Mathematical Analysis*. New York, 1976.

[22] Rienhard Sippel. An experiment on the pure theory of consumer's behaviour. *Economic Journal*, 107:1431–1444, 1997.

[23] Dek Terrell. Incorporating monotonicity and concavity conditions in flexible functional forms. *Journal of Applied Econometrics*, 11:179–194, 1996.

[24] Hal R. Varian. Goodness-of-fit in optimizing models. *Journal of Econometrics*, 46:125–140, 1990.