HETEROGENEITY AND THE NONPARAMETRIC ANALYSIS OF CONSUMER CHOICE: CONDITIONS FOR INVERTIBILITY

Walter Beckert
Richard Blundell

Heterogeneity and the Nonparametric Analysis of Consumer Choice: Conditions for Invertibility

Walter Beckert[*]and Richard Blundell[†]

July 2005

Abstract

This paper considers structural nonparametric random utility models for continuous choice variables. It provides sufficient conditions on random preferences to yield reduced-form systems of nonparametric stochastic demand functions that allow global invertibility between demands and random utility components. Invertibility is essential for global identification of structural consumer demand models, for the existence of well-specified probability models of choice and for the nonparametric analysis of revealed stochastic preference.

Keywords: nonparametric random utility model, stochastic demand, global invertibility.

JEL classification: C14, C31, C51, D1

---

[*]Birkbeck College, University of London, and Institute for Fiscal Studies; corresponding author; address for correspondence: Department of Economics, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK; e-mail: wbeckert@econ.bbk.ac.uk

[†]University College London and Institute for Fiscal Studies

1 Introduction

It is well known from the work of Brown and Walker (1989) that maximization of random utility implies that the heterogeneity components in stochastic demand equations cannot generally be additive, as typically assumed in statistical demand models. Consequently, for recovering consumer preferences across individuals, models should in general be nonseparable in unobserved heterogeneity. However, in nonseparable models identification is typically accomplished through a monotonicity assumption, see for example, Matzkin (2003) and Newey and Imbens (2002). For a multiple equation case, as in consumer choices over several goods, this monotonicity assumption becomes a global invertibility condition. Indeed Matzkin (2005) uses the global invertibility condition to show global identification for general nonparametric simultaneous equation systems. We ask: what conditions on heterogeneous preferences enable such a global invertibility assumption?

One interpretation of randomness in consumer choice data, given prices and total expenditure, is unobserved heterogeneity in consumer preferences. Unobserved preference heterogeneity can be modelled in terms of random utility $U(\mathbf{x}, \epsilon)$ where $\mathbf{x} \in \mathbb{R}_+^J$ is a vector of continuous consumption amounts of J goods and $\epsilon \in \mathbb{R}^{J-1}$ is a $J-1$ dimensional vector representing unobserved heterogeneity in preferences. Then, given prices $\mathbf{p} \in \mathbb{R}_{++}^{J-1}, p_J \equiv 1$, and total expenditure $m > 0$, stochastic demand functions $h(\mathbf{p}, m, \epsilon)$ for the $J-1$ *inside* goods $\mathbf{x}_{-J} = (x_1, \ldots x_{J-1})'$ solve

$$
\begin{aligned}
\mathbf{p} &= \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon) \\
\mathbf{x}_{-J} &= h(\mathbf{p}, m, \epsilon),
\end{aligned}
$$

where $\mathbf{MRS}(\mathbf{x}, \epsilon) = \left[\frac{\partial}{\partial x_j} U(\mathbf{x}, \epsilon) / \frac{\partial}{\partial x_J} U(\mathbf{x}, \epsilon) \right]_{j=1,\ldots,J-1}$ is the $J-1$ dimensional vector of stochastic marginal rates of substitution. We focus on two specific cases. In the first, unobserved heterogeneity is separable in the marginal rate of substitution function $\mathbf{MRS}(\mathbf{x}, \epsilon)$ as in Brown and Matzkin (1995); in the second, this separability is relaxed.

Maximization of random utility implies that the conditional residuals $\nu(\mathbf{p}, m, \epsilon) = h(\mathbf{p}, m, \epsilon) - E[h(\mathbf{p}, m, \epsilon) | \mathbf{p}, m]$ must be functionally dependent on $\mathbf{p}$ and m, so that the heterogeneity components ϵ generally cannot be isolated additively. Lewbel (2001) provides conditions on the reduced form demand system that are necessary and sufficient for statistical demands to satisfy revealed preference inequalities implied by utility maximization. This paper goes beyond

additivity of heterogeneity terms and considers the general consumer choice problem in which unobserved preference heterogeneity is nonseparable. It examines conditions on the structural model $U(x, \epsilon)$ or $\mathbf{MRS}(\mathbf{x}, \epsilon)$ that induce the mapping h between demands $\mathbf{x}_{-J}$ and unobserved preference heterogeneity ϵ, given $\mathbf{p}$ and m, to be globally invertible. In other words, the conditions that imply a continuous one-to-one relationship - a global homeomorphism.

For the special case of scalar heterogeneity $\epsilon \in \mathbb{R}$, global invertibility follows from strict monotonicity of h with respect to ϵ. This paper treats the general case of multidimensional heterogeneity.[1] The global homeomorphism property is necessary for global nonparametric identification of $U(\mathbf{x}, \epsilon)$ (Brown and Matzkin (1995), following the approach taken by Brown (1983) and Roehrig (1988)[2]). It is also required for the existence of well-specified probability models for choice variables $\mathbf{x}_{-J}$, given $\mathbf{p}$ and m, and, hence, for the analysis of revealed stochastic preference (McFadden and Richter (1971, 1990) and McFadden (2004)). In Appendix A we present an example of a deficient probability model in which there are continuous choice variables but they do not have a joint density. In the absence of a proper probability model the postulates of revealed stochastic preference cannot be verified.

The paper proceeds as follows. Section 2 lays out the formal framework and notation for this analysis. Section 3 presents a result for local invertibility of demand functions, primarily as a reference point for the main parts of the paper. Section 4 presents a result on global invertibility when unobserved preference heterogeneity enters the structural model in a separable fashion. This result extends the model considered by Brown and Matzkin (1995). Section 5 discusses this result, illuminating the implicit limitations of the underlying model assumptions in light of a synopsis of further global invertibility results for structural models with completely nonseparable heterogeneity. Section 6 concludes.

---

[1] Brown and Wegkamp (2002) consider the multidimensional case, but their analysis is based on the invertibility assumption I.1 in their paper.

[2] Benkard and Berry (2004) point out that these results are deficient. Recent work by Matzkin (2005) demonstrates how these deficiencies can be remedied and provides a complete characterization of the identification conditions.

2 Framework for Analysis

The analysis in this paper proceeds within the following setup. Let $(\mathbf{X}, \mathcal{X})$ be a metric space of choice variables, where $\mathbf{X} \subset \mathbb{R}^J$ and $\mathcal{X}$ is the Borel σ-algebra of subsets of $\mathbf{X}$.

Denote by $(\mathbf{U}, \mathcal{U}, P_\epsilon)$ the probability space defined over all random (direct) utility functions $U : \mathbb{R}^J_+ \times \mathbb{R}^{J-1} \to \mathbb{R}$, i.e. $U(\mathbf{x}, \epsilon)$, where $\mathbf{x} \in \mathbb{R}^J_+$ is a vector of continuous consumption amounts, $\epsilon \in \mathbb{R}^{J-1}$ is a $J-1$ dimensional random component representing unobserved preference heterogeneity, distributed according to probability measure P_ϵ. $\mathcal{U}$ is the Borel σ-algebra of subsets of $\mathbf{U}$. Elements $U \in \mathbf{U}$ in this probability space satisfy the following assumptions:

Assumption A1: For each ϵ, $U \in \mathcal{U}$ is continuous in its arguments, continuously differentiable in $\epsilon, \mathbf{x}$, strongly monotone, concave and strictly quasi-concave in $\mathbf{x}$.

Assumption A2: Let $\mathbf{MRS}(\mathbf{x}, \epsilon) = \left[\frac{\partial}{\partial x_j} U(\mathbf{x}, \epsilon) / \frac{\partial}{\partial x_J} U(\mathbf{x}, \epsilon) \right]_{j=1,\dots,J-1}$. Suppose that the $(J-1) \times (J-1)$ matrix $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ has full rank $J-1$ for all ϵ.

Assumption A3: (*smoothness in the sense of Debreu*) The bordered Hessian satisfies

$$
\left| \begin{array}{cc} \nabla_{\mathbf{ww}'} U(\mathbf{x}, \epsilon) & \nabla_{\mathbf{w}} U(\mathbf{x}, \epsilon) \\ \nabla_{\mathbf{w}'} U(\mathbf{x}, \epsilon) & 0 \end{array} \right| \neq 0
$$

for all $\mathbf{w}' = (\mathbf{x}', \epsilon')$.

Assumptions A1 - A3 guarantee that the reduced form system of stochastic demands $h(\mathbf{p}, m, \epsilon)$ is a system of continuously differentiable demand functions. In other words, under these assumptions, the system

$$
g(\mathbf{x}_{-J}, m, \mathbf{p}, \epsilon) = \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon) - \mathbf{p}
$$

associates a unique value of $\mathbf{x}_{-J}$ with any $\mathbf{p}$, m and ϵ, i.e. it has a well-defined reduced form $\mathbf{x}_{-J} = h(\mathbf{p}, m, \epsilon)$.

Let $(\mathbf{h}, \mathcal{H}, P_h)$ denote the probability space of demands, where P_h is the probability measure induced by P_ϵ through the nonlinear transformation $h(\mathbf{p}, m, \epsilon)$, given $\mathbf{p}$ and m; and let $\mathcal{H}$ be the Borel σ-field of subsets of h. In the terminology of revealed stochastic preference (McFadden (2004)), the probability spaces $(\mathbf{U}, \mathcal{U}, P_\epsilon)$ and $(\mathbf{h}, \mathcal{H}, P_h)$ are consistent (or $\mathbf{h}$ is $\mathbf{U}$-rational), if

i.a. for any $\mathbf{x}_{-J}$ satisfying $\mathbf{p}'\mathbf{x} \leq m$, $x_J \geq 0$, the inverse image of $\mathbf{x}_{-J} = h(\mathbf{p}, m, \epsilon)$ with respect to ϵ, given $\mathbf{p}$ and m, is in $\mathcal{U}$, i.e. $P_h(h(\mathbf{p}, m, \epsilon)) = P_\epsilon(\tilde{U}(\mathbf{p}, m, \mathbf{x}_{-J}))$, where $\tilde{U}(\mathbf{p}, m, \mathbf{x}_{-J})) = \{U \in \mathbf{U} : (\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J})' = (h(\mathbf{p}, m, \epsilon), x_J)' = \arg\max_{\mathbf{p}'\mathbf{x} \leq m} U(\mathbf{x}, \epsilon)\} \in \mathcal{U}$. In order for unambiguous revelation of stochastic preferences from stochastic demands, this inverse should be unique. This paper provides conditions on the structural model $U \in \mathbf{U}$ that are sufficient for $\mathbf{x}_{-J} = h(\mathbf{p}, m, \epsilon)$ and ϵ to be continuous and globally one-to-one - or a global homeomorphism -, given any $\mathbf{p}$ and m.

It is worth noting that arguments establishing global homeomorphisms rest on applications of the theorems by Gale and Nikaido (1965) or Mas-Colell (1979). These theorems provide sufficient conditions for the existence of global homeomorphisms. Hence, within the constraints of these theorems, there is no scope to determine necessary conditions for global homeomorphisms. To distinguish the conditions for global invertibility of $h(\mathbf{p}, m, \epsilon)$ from the substantially weaker requirements for local invertibility, the following section presents a local invertibility result, while subsequent sections are devoted to conditions for global invertibility. Local invertibility is necessary for global invertibility, and hence the analysis of local invertibility sheds some light on necessary conditions for global invertibility.

3   Local Invertibility

Definition: The random variable $\mathbf{x} \in \mathbb{R}^J$ has dimension J, denoted by $\dim(\mathbf{x}) = J$, if it has a non-degenerate distribution on $\mathbb{R}^J$.

Assumption A4: $\dim(\epsilon) = J - 1$.

Assumption 5: $\mathrm{Im}\,(\mathbf{MRS}(\mathbf{x}, \epsilon)) = \{\mathbf{MRS}(\mathbf{x}, \epsilon) : \epsilon \in \mathbb{R}^{J-1}\} = \mathbb{R}_{++}^{J-1}$.

Lemma 3.1: (*Local Invertibility*) *Suppose A1-A5 hold. Consider the system of demand functions for the $J-1$ inside goods $\mathbf{x}_{-J} = h(\mathbf{p}, m, \epsilon)$. Fix $\epsilon_0 \in \mathbf{R}^{J-1}$. Then, there exists $\delta > 0$ such that, on the sets*

$$\mathcal{E}(\epsilon_0; \delta) \;\; := \;\; \{\epsilon \in \mathbf{R}^{J-1} : ||\epsilon - \epsilon_0|| < \delta\}$$

$$\mathcal{X}(\epsilon_0; \delta) \;\; := \;\; \{\mathbf{z} \in \mathbb{R}_+^{J-1} : \mathbf{z} = h(\mathbf{p}, m, \epsilon) \text{ for } \epsilon \in \mathcal{E}(\epsilon_0; \delta)\}.$$

$\mathbf{x}_{-J} = h(\mathbf{p}, m, \epsilon)$ and ϵ are one-to-one, given any $(\mathbf{p}', m)' > \mathbf{0}$, and hence the distribution of

$x_{-J} \in \mathcal{X}(\epsilon_0; \delta)$, *conditional on* $\mathbf{p}$ *and* m, *is non-degenerate.*

The proof is an application of the Implicit Function Theorem and is omitted. Local invertibility is not enough for global identification of unobserved preference heterogeneity and hence of random utility. The result does, however, point to a necessary condition for global invertibility. Suppose that $h(\mathbf{p}, m, \epsilon)$ is invertible with respect to $\mathbf{p}$. Denote this inverse by $q(\mathbf{x}_{-J}, m, \epsilon)$. The mapping between $\mathbf{x}_{-J}$ and ϵ being homeomorphic, given $\mathbf{p}$ and m, is equivalent to $\mathbf{p} - q(\mathbf{x}_{-J}, m, \epsilon) = \mathbf{0}$ being an implicit homeomorphism between $\mathbf{x}_{-J}$ and ϵ, given $\mathbf{p}$ and m. Let $B_{-J}(\mathbf{p}, m) = \{\mathbf{x}_{-J} \in \mathbb{R}_+^{J-1} : \mathbf{p}'\mathbf{x}_{-J} + x_J = m, x_J \geq 0\}$. Since

$$\mathbf{p} = q(\mathbf{x}_{-J}, m, \epsilon) = \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon),$$

this implies that $\mathbf{x}_{-J}$ and ϵ are one-to-one, given $\mathbf{p}$ and m, if, and only if, for any $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, $\mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon)$ is an implicit homeomorphism between $\mathbf{x}_{-J}$ and ϵ. Under conventional smoothness assumptions, the rank condition on the matrix $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ in A2, $\mathrm{rk}(\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)) = J - 1$ on $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, is a necessary, though not sufficient condition for this.

4   Global Invertibility with MRS-Separable Heterogeneity

This section examines structural model specifications in which unobserved preference heterogeneity ϵ enters the marginal rate of substitution in a separable form. Specifically, it considers models for marginal rates of substitution in which unobserved preference heterogeneity enters in a multiplicative fashion. The advantage of such specifications is that they permit higher order derivatives of random utility to depend on unobserved heterogeneity as well, allowing i.a. for heterogeneous curvature of utility and heterogenous substitution elasticities. They include the model of Brown and Matzkin (1995) as a special case; in their model, heterogeneity enters random utility linearly, so that the utility curvature is not heterogeneous.

The following additional assumptions are maintained:

Assumption A4': In addition to A4, assume that supp(ϵ) is a rectangle.

This assumption is rather mild, as McFadden and Train (2000) demonstrate. Their Lemma 3 shows that for any random vector $\xi \in \mathbb{R}^{J-1}$ which admits a conditional probability law for each coordinate there exists a uniformly distributed random vector $\epsilon \in [0,1]^{J-1}$ and measurable functions $g_i : [0,1]^i \to \mathbb{R}$, $i = 1, \ldots, J-1$, such that $\xi_i = g_i(\epsilon_1, \ldots, \epsilon_i)$ almost surely.

Assumption A6: $\mathbf{MRS}(\mathbf{x}, \epsilon)$ is multiplicatively separable with respect to ϵ:

$$\mathbf{MRS}(\mathbf{x}, \epsilon) = v(\mathbf{x}) + K(\mathbf{x})\psi(\epsilon),$$

where $v(\mathbf{x})$ is a $(J-1) \times 1$ vector of nonnegative functions, $K(\mathbf{x})$ is a $(J-1) \times (J-1)$ matrix with full rank and span equal to R_{++}^{J-1}, and $\psi : \mathbf{R}^{J-1} \to \mathbf{R}^{J-1}$ satisfies the Gale and Nikaido or Mas-Colell conditions.[3]

Lemma 4.1: *Suppose A1, A2, A3, A4' and A6 hold. Then, for any $\mathbf{p}$ and m, $h(\mathbf{p}, m, \epsilon)$ is globally invertible for all $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, and, hence, $\mathbf{x}_{-J}$ has a non-degenerate distribution on $B_{-J}(\mathbf{p}, m)$, given any $\mathbf{p}$ and m.*

Proof: From the first-order conditions and A8,

$$\psi(\epsilon) = K(\mathbf{x})^{-1}(\mathbf{p} - v(\mathbf{x})),$$

and the result follows from an application of the Gale Nikaido or the Mas-Colell Theorem. $\square$

As an illustration of the these results, consider the random utility model

$$U(\mathbf{x}, \epsilon) = u(\mathbf{x}_{-J})'\psi(\epsilon) + \nu(\mathbf{x}),$$

where $u(\cdot)$ is defined on $\mathbb{R}_+^{J-1}$, monotonically increasing and weakly concave, $\nu(\cdot)$ is defined on $\mathbb{R}_+^J$ and satisfies A1 and A3, and $\psi(\epsilon)$ as in A4'. In this model, preferences are nonseparable over the J goods, and marginal utilities may involve any subset of the components of ϵ. Then,

$$
\begin{aligned}
\mathbf{MRS}(\mathbf{x}, \epsilon) &= \left[\frac{\frac{\partial}{\partial x_j}\nu(\mathbf{x})}{\frac{\partial}{\partial x_J}\nu(\mathbf{x})} \right]_{j=1,\ldots,J-1} + \left[\frac{\partial}{\partial x_J}\nu(\mathbf{x}) \right]^{-1} \left[\frac{\partial}{\partial x_j} u(\mathbf{x}_{-J})' \right]_{j=1,\ldots,J-1} \psi(\epsilon) \\
&= v(\mathbf{x}) + K(x)\psi(\epsilon),
\end{aligned}
$$

---

[3]The Gale and Nikaido conditions are: the support of ϵ is a rectangle, and the Jacobian of $\psi(\epsilon)$ is a P matrix for every ϵ, i.e. every principal minor has positive sign. The Mas-Colell conditions are weaker and imply that, if the support of ϵ is a rectangle, the Jacobian needs to be a P matrix only at its vertices and that for ϵ in the interior of its support it is only required that the Jacobian have a positive determinant. An example of a function $\psi(\epsilon)$ is provided in Appendix B.

where $v(\mathbf{x}) = \left[\dfrac{\frac{\partial}{\partial x_j} \nu(\mathbf{x})}{\frac{\partial}{\partial x_J} \nu(\mathbf{x})} \right]_{j=1,\ldots,J-1} \in \mathbb{R}_+^{J-1}$ and $K(\mathbf{x}) = \left[\frac{\partial}{\partial x_J} \nu(\mathbf{x}) \right]^{-1} \left[\frac{\partial}{\partial x_j} u(\mathbf{x}_{-J})' \right]_{j=1,\ldots,J-1}$. A1 and A3 imply that $K(\mathbf{x})$ has full rank and that its span is $\mathbb{R}_{++}^{J-1}$.

The model due to Brown and Matzkin (1995) can be obtained by choosing $u(\cdot)$ and $\psi(\cdot)$ the respective identity functions, i.e. $u(\mathbf{x}_{-J}) = \mathbf{x}_{-J}$ for any $\mathbf{x}_{-J} \in \mathbb{R}_+^{J-1}$, and $\psi(\epsilon) = \epsilon$ for any ϵ, and $\nu(\mathbf{x}) = \phi(\mathbf{x}) + x_J$, so that $U(\mathbf{x}, \epsilon) = \phi(\mathbf{x}) + \mathbf{x}'_{-J} \epsilon + x_J$. Brown and Matzkin's model implies that marginal rates of substitution are additive in ϵ, hence invertibility follows directly from the first-order conditions and no recourse to the Gale Nikaido or Mas-Colell results is necessary, so that ϵ need not have rectangular support. Another illustration is provided by a random coefficient Cobb-Douglas utility model, where the random coefficients are functions of ϵ satisfying Gale Nikaido or Mas-Colell conditions.

5 Global Invertibility without MRS-Separable Heterogeneity

Section 4 extends the class of nonparametric random utility models which are globally identifiable beyond the model of Brown and Matzkin (1995). To illuminate the restrictions that the separability assumption A6 imposes, this section considers a variety of structural models where the marginal rate of substitution is not separable in heterogeneity. Such cases require stronger assumptions. The reason is that the property of $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ being a P matrix, in general, must be deduced from

$$\nabla_\epsilon h(\mathbf{p}, m, \epsilon) = - \left[\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon) \right]^{-1} \nabla_\epsilon MRS(\mathbf{x}, \epsilon)$$

where $\mathbf{x} = (h(\mathbf{p}, m, \epsilon)', m - h(\mathbf{p}, m, \epsilon)'\mathbf{p})'$, for any $\mathbf{p}, m, \epsilon$. Hence, assumptions on the structural preference model, i.e. on $\mathbf{MRS}(\mathbf{x}, \epsilon)$, need to be such that, after differentiation with respect to $\mathbf{x}_{-J}$ and ϵ, inversion of $\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon)$ and multiplication of $- \left[\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon) \right]^{-1}$ and $\nabla_\epsilon MRS(\mathbf{x}, \epsilon)$, the resulting matrix has principal minors of the same sign. This section provides different sets of assumptions that induce this property.

Assumption A7: For $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, $U(\mathbf{x}, \epsilon)$ satisfies $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon) \nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon) = A(\mathbf{x}, \epsilon)$ negative definite a.s., and $\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon)$ full rank and symmetric, a.s.

Lemma 5.1: *Suppose that A1, A3, A4', and A7 hold. Then, for any $\mathbf{p}$ and m, $h(\mathbf{p}, m, \epsilon)$ is globally invertible for all $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, and, hence, $\mathbf{x}_{-J}$ has a non-degenerate distribution on $B_{-J}(\mathbf{p}, m)$, given any $\mathbf{p}$ and m.*

Proof: By A3, $\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon)$ has full rank, so that its inverse exists. From the first-order conditions,

$$\nabla_\epsilon h(\mathbf{p}, m, \epsilon) = -\left[\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon)\right]^{-1} \nabla_\epsilon \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon).$$

The Gale Nikaido Theorem requires that this $(J-1) \times (J-1)$ Jacobian matrix has all principal minors positive. Magnus and Neudecker, Chapt.1 Theorem 29, establishes that for symmetric matrices this is equivalent to it being positive definite. Therefore, A7 implies

$$\nabla_\epsilon \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon) = A(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon) \left[\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \epsilon)\right]^{-1},$$

so that $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ is seen to be positive definite. $\square$

The following corollary follows immediately.

Corollary: *Suppose that A1, A3, A4', and A7 hold. Then, $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ is positive definite symmetric for all $\mathbf{p}, m, \epsilon$.*

(A constructed) Example: Suppose that

$$U(x_1, x_2, x_3) = \left(\alpha\epsilon_1 \exp(\frac{x_1}{\epsilon_1}) + \exp(x_3)\right)^{\epsilon_1} + \left(\beta\epsilon_2 \exp(\frac{x_2}{\epsilon_2}) + \exp(x_3)\right)^{\epsilon_2},$$

where α, β are positive parameters and ϵ_1, ϵ_2 are random components; $\epsilon_i > 0$ is necessary and sufficient for strict monotonicity and strict quasi-concavity, and $\epsilon_i < 1$ is required for concavity, $i = 1, 2$. Here, (x_1, x_3) and (x_2, x_3) are nonseparable. Moreover,

$$\mathbf{MRS}(\mathbf{x}, \epsilon) = \begin{bmatrix} \alpha \exp(\frac{x_1}{\epsilon_1} - x_3) \\ \beta \exp(\frac{x_2}{\epsilon_2} - x_3) \end{bmatrix}$$

shows that the model is nonseparable in the stochastic components. Since

$$\nabla_{\mathbf{x}_{-3}} \mathbf{MRS}(\mathbf{x}, \epsilon) = \begin{bmatrix} \frac{\alpha}{\epsilon_1} \exp(\frac{x_1}{\epsilon_1} - x_3) & 0 \\ 0 & \frac{\beta}{\epsilon_2} \exp(\frac{x_2}{\epsilon_2} - x_3) \end{bmatrix}$$

$$\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon) = \begin{bmatrix} -\frac{\alpha}{\epsilon_1^2} \exp(\frac{x_1}{\epsilon_1} - x_3) & 0 \\ 0 & -\frac{\beta}{\epsilon_2^2} \exp(\frac{x_2}{\epsilon_2} - x_3) \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{x_1}{\epsilon_1} & 0 \\ 0 & -\frac{x_2}{\epsilon_2} \end{bmatrix} \nabla_{\mathbf{x}_{-3}} \mathbf{MRS}(\mathbf{x}, \epsilon),$$

9

it follows that

$$\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon) \nabla_{\mathbf{x}_{-3}} \mathbf{MRS}(\mathbf{x}, \epsilon) = \begin{bmatrix} -\frac{x_1}{\epsilon_1} & 0 \\ 0 & -\frac{x_2}{\epsilon_2} \end{bmatrix} \left[\nabla_{\mathbf{x}_{-3}} \mathbf{MRS}(\mathbf{x}, \epsilon) \right]^2,$$

a diagonal matrix with negative elements on the diagonal, almost surely. Hence A7 is met.

An assumption slightly weaker than A7 is[4]

Assumption A8: $U(\mathbf{x}, \epsilon)$ is strictly concave in $\mathbf{x}_{-J}$ and linear in the outside good x_J, and $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ is positive definite almost surely.

Lemma 5.2: *Suppose that A1, A3, A4', and A8 hold. Then, for any $\mathbf{p}$ and m, $h(\mathbf{p}, m, \epsilon)$ is globally invertible for all $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, and, hence, $\mathbf{x}_{-J}$ has a non-degenerate distribution on $B_{-J}(\mathbf{p}, m)$, given any $\mathbf{p}$ and m.*

Proof: A8 implies that $-\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon)$ is positive definite for all $\mathbf{x}$ and ϵ, and symmetric. Its inverse inherits these properties. Horn and Johnson, Theorem 7.6.3, then implies that its product with a positive definite matrix $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ is diagonalizable, i.e. similar[5] to a diagonal matrix, whose eigenvalues are positive. Similarity means that there exists a nonsingular transformation S of $x_{-J} = h(\mathbf{p}, m, \epsilon)$, possibly dependent on $\mathbf{p}, m, \epsilon$, such that the transformed vector of demands has a distribution, conditional on $\mathbf{p}$ and m, that can be deduced from the distribution of ϵ by evaluation at the inverse function and multiplication by a Jacobian which is diagonal. Then, one diagonalization is

$$\nabla_\epsilon h(\mathbf{p}, m, \epsilon) = S(\mathbf{p}, m, \epsilon) D(\mathbf{p}, m, \epsilon) S(\mathbf{p}, m, \epsilon)^{-1},$$

where $S(\mathbf{p}, m, \epsilon)$ is a nonsingular matrix consisting of the $J-1$ eigenvectors of $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ and $D(\mathbf{p}, m, \epsilon)$ is a diagonal matrix with the positive eigenvalues of $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ on its diagonal. This is necessary and sufficient for $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ to be positive definite almost surely. Note that under A10 $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ is not necessarily symmetric, so that the Magnus and Neudecker result cannot be applied and the Gale Nikaido conditions need to be verified. For $k = 1, \ldots, J-1$, define $k \times (J-1)$ matrices $E_k = [\mathbf{I}_k, \mathbf{0}]$, where $\mathbf{0}$ is a $(J-1-k) \times (J-1)$ matrix of zeros. Then, the kth principal minor of the Jacobian matrix $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ is

$$|\nabla_\epsilon h(\mathbf{p}, m, \epsilon)_k| = |E_k \nabla_\epsilon h(\mathbf{p}, m, \epsilon) E_k'| = |E_k S(\mathbf{p}, m, \epsilon) D(\mathbf{p}, m, \epsilon) S(\mathbf{p}, m, \epsilon)^{-1} E_k|.$$

---

[4]A8 appears to be weaker than A9 because it does not imply symmetry of the Jacobian $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$.

[5]An $n \times n$ matrix A is similar to an $n \times n$ matrix B if there exists a nonsingular $n \times n$ matrix S such that $B = S^{-1} AS$. Similarity is an equivalence relation. See Horn and Johnson, section 1.3, for further details.

Therefore, for any $\mathbf{y} \in \mathbf{R}^k, \mathbf{y} \neq \mathbf{0}$, and any $k = 1, \ldots, J-1$,

$$
\begin{aligned}
\mathbf{y}' \nabla_\epsilon h(\mathbf{p}, m, \epsilon)_k \mathbf{y} &= \mathbf{y}' E_k \nabla_\epsilon h(\mathbf{p}, m, \epsilon) E_k' \mathbf{y} \\
&= (E_k' \mathbf{y})' \nabla_\epsilon h(\mathbf{p}, m, \epsilon)(E_k' \mathbf{y}) > 0,
\end{aligned}
$$

where $E_k' \mathbf{y} \neq 0$ and the last inequality follows because $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ is positive definite almost surely. Hence, the Jacobian has all principal submatrices positive definite almost surely. Therefore, for any $k = 1, \ldots, J-1$, there exists a full-rank $k \times k$ matrix $P_k(\mathbf{p}, m, \epsilon)$ such that

$$
\begin{aligned}
\nabla_\epsilon h(\mathbf{p}, m, \epsilon)_k &= P_k(\mathbf{p}, m, \epsilon) P_k(\mathbf{p}, m, \epsilon)' \\
\Rightarrow |\nabla_\epsilon h(\mathbf{p}, m, \epsilon)_k| &= |P_k(\mathbf{p}, m, \epsilon) P_k(\mathbf{p}, m, \epsilon)'| \\
&= |P_k(\mathbf{p}, m, \epsilon)|^2 > 0.
\end{aligned}
$$

Therefore, the Gale Nikaido conditions are satisfied. $\qquad\square$

Quasi-linearity $U(\mathbf{x}, \epsilon)$ in the outside good x_J leaves the possibility that the demand for the outside good may be negative. It can be relaxed if, instead, more structure is imposed on $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ and $\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon)$. Consider, for instance,

Assumption A9: (i) $\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon)$ has negative diagonal and non-negative off-diagonal entries, a.s.; (ii) $(-1)^J \nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ has positive diagonal and non-positive off-diagonal entries, and all its principal minors are positive, a.s.; and (iii) $(-1)^J \nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon) - \nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon) \geq \mathbf{0}$, a.s.

Matrices, having properties as in (i) and (ii), are sometimes referred to as M-matrices; see, e.g., Horn and Johnson (1991). Assumption A9 is particularly attractive because it has an economic interpretation. It can be shown that the restriction imposed by A9 imply that the inside goods $x_1, \ldots, x_{J-1}$ are pairwise symmetric gross substitutes, while x_j and x_J, $j = 1, \ldots, J-1$, are pairwise not necessarily symmetric gross complements.[6] Let $\mathbf{Z}$ be the class of square matrices whose off-diagonal elements are all non-positive, as in Fiedler and Pták's definition (4,1). And let $\mathbf{K}$ be those elements in $\mathbf{Z}$ which have all principal minors positive,

---

[6]Illustrative examples are available from the authors upon request. With additional assumptions on the principal minors, such sign restrictions imply, furthermore, that $\nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ and $- \left[\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \epsilon) \right]^{-1}$ are strictly totally positive and bounded almost surely. Applying the Cauchy-Binet formula to $|h(\mathbf{p}, m, \epsilon)_k|$, $k = 1, \ldots, J-1$, and using a result due to Karlin (1968) on totally positive matrices (chapt.5, Theorem 3.1), it can be shown immediately that $\nabla_\epsilon h(\mathbf{p}, m, \epsilon)$ is a P matrix.

as in Fiedler and Pták's definition (4,4). Lemma 5.4 below uses Fiedler and Pták's Theorem (4,6): If $A \in \mathbf{K}$, $B \in \mathbf{Z}$ and $B - A \geq \mathbf{0}$, then, i.a., $B^{-1}A \in \mathbf{K}$.

Lemma 5.3: *Suppose that A1, A3, A4', and A9 hold. Then, for any* $\mathbf{p}$ *and* m, $h(\mathbf{p}, m, \epsilon)$ *is globally invertible for all* $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, *and, hence,* $\mathbf{x}_{-J}$ *has a non-degenerate distribution on* $B_{-J}(\mathbf{p}, m)$, *given any* $\mathbf{p}$ *and* m.

Proof: By A10(i), $-\nabla_{\mathbf{x}_{-J}}\mathbf{MRS}(\mathbf{x}, \epsilon)$ has positive diagonal and non-positive off-diagonal entries. Hence it belongs to the class $\mathbf{Z}$; and by A10(ii), $(-1)^J \nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon)$ belongs to the class $\mathbf{K}$. Hence, using A10(iii), by Fiedler and Pták, Theorem (4,6),

$$
\begin{aligned}
h(\mathbf{p}, m, \epsilon) &= -\left[\nabla_{\mathbf{x}_{-J}}\mathbf{MRS}(\mathbf{x}, \epsilon)\right]^{-1} \nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon) \\
&= (-1)^J \left[-\nabla_{\mathbf{x}_{-J}}\mathbf{MRS}(\mathbf{x}, \epsilon)\right]^{-1} \nabla_\epsilon \mathbf{MRS}(\mathbf{x}, \epsilon) \in \mathbf{K},
\end{aligned}
$$

i.e. all its principal minors are positive, so that the Gale Nikaido conditions are satisfied. $\qquad \square$

Among the structural models with non-separable heterogeneity, models based on A9, in light of its economic interpretability, are likely to be the most useful for applied work.

6 Conclusions

This paper provides conditions on structural nonparametric preference models for continuous choices under which the induced stochastic demand system is non-separable in unobserved preference heterogeneity and globally invertible. It extends the class of nonparametrically identifiable random utility models with separable heterogeneity beyond the classical model of Brown and Matzkin (1995) and discusses various extensions to completely non-separable cases. This broadens the class of random utility models suitable for nonparametric microeconometric analysis.

The synopsis of these conditions emphasizes the view that microeconometric modelling of demand acknowledging unobserved preference heterogeneity requires additional restrictions on preferences, beyond those imposed by microeconomic theory.

A   Example: Non-existence of joint density

Consider the following random utility model: $U(x_1, x_2, \epsilon) = \min\{x_1 + x_2, \epsilon x_1 + \frac{1}{2}x_2\}$, where $\text{supp}(\epsilon) = (\frac{1}{2}, +\infty)$. Indifference curves associated with this random utility model are kinked, and the location and angle of the kink are determined by ϵ. Depending on relative prices p_{x_1}/p_{x_2} and given any income m, various types of solutions to the consumer's utility maximization problem can arise. Interior solutions for $p_{x_1}/p_{x_2} \in (1, 2\epsilon)$ are characterized by $\frac{\bar{x}_2^\star}{\bar{x}_1^\star} = 2(\epsilon - 1)$ and yield $\bar{x}_i^\star = (2(\epsilon - 1))^{i-1} m/(p_{x_1} + 2(\epsilon - 1)p_{x_2})$, $i = 1, 2$; set-valued solutions arise when either $p_{x_1}/p_{x_2} = 1$, in which case $x_1^\star + x_2^\star = u^\star$, where $u^\star = m/(p_{x_1} + p_{x_2})$ and $x_1^\star \in [\bar{x}_1^\star, m/p_{x_1}]$, $x_2^\star \in [0, \bar{x}_2^\star]$; or when $p_{x_1}/p_{x_2} = 2\epsilon$, in which case $\epsilon x_1^\star + \frac{1}{2}x_2^\star = u^\star = m/2p_{x_2}$ and $x_1^\star \in [0, \bar{x}_1^\star]$, $x_2^\star \in [\bar{x}_2^\star, m/p_{x_2}]$; corner solutions arise when $p_{x_1}/p_{x_2} < 1$ or $p_{x_1}/p_{x_2} > 2\epsilon$.

Now suppose that for a consumer with income $m = 27$, consumption choices $x_1 = 3$ and $x_2 = 18$ are observed at prices $p_{x_1} = 3, p_{x_2} = 1$. Assuming the consumer maximizes $U(x_1, x_2, \epsilon)$, this could either be a corner solution, in which case one infers $\epsilon_1 = 4$; or it could be an element of a set-valued solution, in which case one infers $\epsilon_2 = 3/2$. This amounts to a lack of identification of the structural model. If, in the spirit of revealed preference type comparisons, the price of good one changes to $p_{x_1} = 1$, then ϵ_1 induces another solution in the set $x_1 \in [3, 27]$ and $x_2 = 27 - x_1$, while this ϵ_2 induces a solution in the larger set $x_1 \in [27/2, 27]$ and $x_2 = 27 - x1$. Note that, in fact, given $p_{x_1} = p_{x_2} = 1$ any choice pair $\{(x_1, x_2) : x_1 \leq 27/2, x_1 + x_2 = 27\}$ can be induced by a continuum of values of ϵ, namely all $\epsilon \geq \frac{1}{2}\frac{x_2}{x_1} + 1 \geq \frac{3}{2}$. This implies that any such (x_1, x_2) is observed with positive probability induced by ϵ, $\Pr(\epsilon \geq \frac{1}{2}\frac{x_2}{x_1} + 1)$. This is a deficient probability probability model, since x_1 and x_2 are continuous choice variables, but do not have a joint density.

B   Example: Function $\psi(\cdot)$ satisfying Gale Nikaido and Mas-Colell Conditions

Consider the function $\psi(\epsilon) = \exp(\mathbf{A}\epsilon)$, where $\epsilon \in \mathcal{E} \subset \mathbb{R}^{J-1}$ and $\mathbf{A}$ is a $(J-1) \times (J-1)$ matrix. The Jacobian of $\psi(\epsilon)$ is

$$\nabla_\epsilon \psi(\epsilon) = [\psi(\epsilon) \circ \mathbf{A}_{1.}, \ldots, \psi(\epsilon) \circ \mathbf{A}_{J-1.}],$$

where $\mathbf{A}_{j\cdot}$, $j = 1, \ldots, J - 1$, is the jth column of $\mathbf{A}$.

Suppose $\mathbf{A}$ is triangular, with positive diagonal elements. Then, $\nabla_\epsilon \psi(\epsilon)$ is triangular as well and, since $\psi(\epsilon)$ has positive elements a.s., has positive diagonal elements, and the same is true for every principal submatrix of $\mathbf{A}$ and $\nabla_\epsilon \psi(\epsilon)$. Hence, $\nabla_\epsilon \psi(\epsilon)$ and its principal submatrices have determinants which equal their traces and hence are positive. Consequently, $\psi(\epsilon)$ also satisfies the weaker Mas-Colell conditions.

References

[1] Brown, B.W. (1983): "The Identification Problem in Systems Nonlinear in the Variables", *Econometrica*, **51**(1), 175-196

[2] Brown, B.W. and M.B. Walker (1989): "The Random Utility Hypothesis and Inference in Demand Systems", *Econometrica*, **57**(4), 815-829

[3] Brown, D.J. and R.L. Matzkin (1995): "Estimation of a Random Utility Model from Data on Consumer Demand", mimeo, Yale University and Northwestern University

[4] Brown, D.J. and M.H. Wegkamp (2002): "Weighted Minimum Mean-Square Distance from Independence Estimation", *Econometrica*, **70**(5), 2035-2051

[5] Benkard, C.L. and S. Berry (2004): "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on B. Brown (1983) and Roehrig (1988)", mimeo, Stanford University and Yale University

[6] Fiedler, M. and V. Pták (1962): "On Matrices with Non-Positive Off-Diagonal Elements and Positive Principal Minors", *Czechoslovak Mathematical Journal*, **12**, 382-400

[7] Gale, D. and H. Nikaido (1965): "The Jacobian matrix and global univalence of mappings", *Mathemathische Annalen*, **159**(2), 81-93

[8] Horn, R.A. and C.R. Johnson (1985): *Matrix Analysis*, Cambridge, MA: MIT Press

[9] Horn, R.A. and C.R. Johnson (1991): *Topics in Matrix Analysis*, Cambridge, UK: Cambridge University Press

[10] Karlin, S. (1968): *Total Positivity*, Stanford: Standford University Press

[11] Lewbel, A. (2001): "Demand Systems with and without Errors", *American Economic Review*, **91**(3), 611-618

[12] Magnus, J.R. and H. Neudecker (1988): *Matrix Differential Calculus with Applications in Statistics and Econometrics*, New York: Wiley

[13] Mas-Colell, A. (1979): "Homeomorphisms of Compact, Convex Sets and the Jacobian Matrix", *SIAM Journal of Mathematical Analysis*, **10**(6), 1105-1109

[14] Matzkin, R.L. (2003): "Nonparametric Estimation of Nonadditive Random Functions", *Econometrica*, **71**(5), 1339-1376

[15] Matzkin, R.L. (2005): "Nonparametric Simulataneous Equations", mimeo, Department of Economics, Northwestern University, June.

[16] McFadden, D.L. (2004): "Revealed Stochastic Preference: A Synthesis", mimeo, UC Berkeley

[17] McFadden, D.L. and K. Richter (1971): "On the Extension of a Set Function on a Set of Events to a Probability on the Generated Boolean σ-algebra", UC Berkeley, working paper

[18] McFadden, D.L. and K. Richter (1990): "Stochastic Rationality and Revealed Stochastic Preference", in J. Chipman, D. McFadden, K. Richter, eds.: *Preference, Uncertainty, and Rationality*, Westview Press, 161-186

[19] McFadden, D.L. and K. Train (2000): "Mixed MNL Models for Discrete Response", *Journal of Applied Econometrics*, **15**(5), 447-470

[20] Newey, W. amd G. Imbens (2002), "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity" draft Department of Economics, M.I.T., October.

[21] Roehrig, C.S. (1988): "Conditions for Identification in Nonparametric and Parametric Models", *Econometrica*, **56**(2), 433-447