THE REVIEW OF ECONOMIC STUDIES

Review of Economic Studies (2008) **75**, 1069–1080 0034-6527/08/00431069$02.00

Heterogeneity and the Non-Parametric Analysis of Consumer Choice: Conditions for Invertibility

WALTER BECKERT
Birkbeck College and Institute for Fiscal Studies

and

RICHARD BLUNDELL
University College London and Institute for Fiscal Studies

First version received August 2005; *final version accepted January* 2008 (*Eds.*)

This paper considers structural non-parametric random utility models for continuous choice variables with unobserved heterogeneity. We provide sufficient conditions on random preferences to yield reduced-form systems of non-parametric stochastic demand functions that allow global invertibility between demands and non-separable unobserved heterogeneity. Invertibility is essential for global identification of structural consumer demand models, for the existence of well-specified probability models of choice and for the non-parametric analysis of revealed stochastic preference. We distinguish between new classes of models in which heterogeneity is separable and non-separable in the marginal rates of substitution, respectively.

1. INTRODUCTION

Heterogeneity in consumer choices is generally recognized as an empirical regularity (*e.g.* Heckman, 2001). A growing econometric literature attempts to model heterogeneous consumer demand within the classical microeconomic consumer choice framework of utility maximization (*e.g.* Brown and Matzkin, 1995; Beckert, 2006; Browning and Carro, 2007; Lewbel, 2007; Matzkin, 2005, 2007). Unobserved heterogeneity is thereby modelled in terms of random utility. It is well known from the work of Brown and Walker (1989) that maximization of random utility implies that the random heterogeneity components in stochastic demand equations cannot generally be additive, as typically assumed in statistical demand models. Consequently, in order to be consistent with random utility maximization, demand models should, in general, be non-separable in unobserved heterogeneity. This, however, imposes further requirements for global econometric identification of utility from demands. In non-separable demand models for a single good identification is typically accomplished through a monotonicity assumption; see, for example, Matzkin (2003) and Imbens and Newey (2002). For the case of consumer choice over several goods, inducing a system of demands, this monotonicity assumption becomes a global invertibility condition. Indeed, Matzkin (2005) uses the global invertibility condition to show global identification for general non-parametric simultaneous equation systems. We ask: what conditions on heterogeneous preferences enable such a global invertibility assumption?

More formally, unobserved preference heterogeneity can be modelled in terms of random utility $U(\mathbf{x},\varepsilon)$, where $\mathbf{x}\in\mathbb{R}^J_+$ is a vector of continuous consumption amounts of J goods and $\varepsilon\in\mathbb{R}^{J-1}$ 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 $m>0$, the system of stochastic demands $h(\mathbf{p},m,\varepsilon)$ for the $J-1$ *inside* goods $\mathbf{x}_{-J}=(x_1,\dots x_{J-1})'$ solves

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

$$\mathbf{x}_{-J}=h(\mathbf{p},m,\varepsilon),$$

where, assuming differentiability, $\mathbf{MRS}(\mathbf{x},\varepsilon)=\left[\frac{\partial}{\partial x_j}U(\mathbf{x},\varepsilon)/\frac{\partial}{\partial x_J}U(\mathbf{x},\varepsilon)\right]_{j=1,\dots,J-1}$ is the $J-1$-dimensional vector of stochastic marginal rates of substitution (MRS). Utility maximization implies that the conditional residuals $\nu(\mathbf{p},m,\varepsilon)=h(\mathbf{p},m,\varepsilon)-E[h(\mathbf{p},m,\varepsilon)|\mathbf{p},m]$ must be functionally dependent on $\mathbf{p}$ and m, so that the heterogeneity components ε generally cannot be isolated additively (Brown and Walker, 1989). 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 ε is non-separable in $h(\mathbf{p},m,\varepsilon)$.

Given ε, utility maximization implies that the system $h(\mathbf{p},m,\varepsilon)$ satisfies integrability conditions, so that the underlying utility function could be recovered if ε were known. Global identification of $U(\mathbf{x},\varepsilon)$ then requires that, given $\mathbf{p}$ and m, a unique ε can be identified with $\mathbf{x}_{-J}=h(\mathbf{p},m,\varepsilon)$. Under assumptions on $U(\mathbf{x},\varepsilon)$ that guarantee that induced demands $h(\mathbf{p},m,\varepsilon)$ form a system of continuous functions, rather than correspondences, this requirement is equivalent to the mapping h between demands $\mathbf{x}_{-J}$ and unobserved preference heterogeneity ε, given $\mathbf{p}$ and m, being globally homeomorphic, that is, continuous and one to one. Global identification under invertibility assumptions is treated by Brown and Matzkin (1995), following the approach taken by Brown (1983) and Roehrig (1988).[1]

The global homeomorphism property 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).[2] In the absence of a proper probability model the postulates of revealed stochastic preference cannot be verified, because for an unambiguous revelation of stochastic preference, it is necessary that stochastic demands be globally invertible. The results provided in this paper guarantee global invertibility.

Unobserved preference heterogeneity ε can enter $U(\mathbf{x},\varepsilon)$ in complex ways. We distinguish two specific cases. In the first, unobserved heterogeneity is multiplicatively separable in the marginal rate of substitution function $\mathbf{MRS}(\mathbf{x},\varepsilon)$, as in Brown and Matzkin (1995); in the second, this separability is relaxed. For these cases, this paper examines conditions on the structural model $U(\mathbf{x},\varepsilon)$ or $\mathbf{MRS}(\mathbf{x},\varepsilon)$ that induce the mapping h between demands $\mathbf{x}_{-J}$ and unobserved preference heterogeneity ε, given $\mathbf{p}$ and m, to be globally homeomorphic.

The paper proceeds as follows. Section 2 lays out the formal framework and notation for this analysis. Section 3 presents a result on global invertibility when unobserved preference heterogeneity enters the structural model in a multiplicatively separable fashion. This result extends the model considered by Brown and Matzkin (1995). Section 4 discusses the general case of non-separable heterogeneity. Section 5 concludes.

1. Benkard and Berry (2006) 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. In an earlier working paper, Beckert and Blundell (2005), 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 with respect to Lebesgue measure. This appendix can be accessed at www.restud.org.uk/supplementary.

2. FRAMEWORK FOR ANALYSIS

The analysis in this paper proceeds within the following set-up. Denote by $(\mathbf{U}, \mathcal{U}, P_U)$ the probability space defined over a set of random (direct) utility functions $U : \mathbb{R}^J_+ \times \mathbb{R}^{J-1} \to \mathbb{R}$, that is, $U(\mathbf{x}, \varepsilon)$. Here, $\mathbf{x} \in \mathbb{R}^J_+$ is a vector of continuous consumption amounts; $\varepsilon \in \mathbb{R}^{J-1}$ is a $J-1$-dimensional random component representing unobserved preference heterogeneity, distributed according to probability measure P_ε, independent of $\mathbf{p}$ and m; P_U is the probability measure on $\mathcal{U}$ induced by P_ε, that is, $P_U(U \in \mathcal{U}) = P_\varepsilon(\varepsilon : U(\cdot, \varepsilon) \in \mathcal{U})$; and $\mathcal{U}$ 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, $U \in \mathbf{U}$ is continuous in its arguments, twice continuously differentiable in $\varepsilon, \mathbf{x}$, strictly monotone, concave, and strictly quasi-concave in $\mathbf{x}$.*

Assumption A2 (Smoothness in the Sense of Debreu). *The bordered Hessian satisfies*

$$\begin{vmatrix} \nabla_{\mathbf{ww}'} U(\mathbf{x}, \varepsilon) & \nabla_{\mathbf{w}} U(\mathbf{x}, \varepsilon) \\ \nabla_{\mathbf{w}'} U(\mathbf{x}, \varepsilon) & 0 \end{vmatrix} \neq 0$$

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

Assumption A1 ensures that the induced system of demands $h(\mathbf{p}, m, \varepsilon)$ constitutes a system of functions, rather than correspondences. Assumption A2 extends the condition on the bordered Hessian beyond $\mathbf{x}$ to ε. This, in turn, extends the induced continuous differentiability of h from $\mathbf{p}$ and m to ε. It should be emphasized that these assumptions suffice to establish classical microeconomic properties of demand functions for their stochastic representations $h(\mathbf{p}, m, \varepsilon)$. Define the vector of marginal rates of substitution as

$$\mathbf{MRS}(\mathbf{x}, \varepsilon) = \left[\frac{\partial}{\partial x_j} U(\mathbf{x}, \varepsilon) \Big/ \frac{\partial}{\partial x_J} U(\mathbf{x}, \varepsilon) \right]_{j=1,\dots,J-1}.$$

Then, under these assumptions, the implicit system

$$\mathbf{0} = g(\mathbf{x}_{-J}, m, \mathbf{p}, \varepsilon) = \mathbf{MRS}(\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J}, \varepsilon) - \mathbf{p}$$

associates a unique value of $\mathbf{x}_{-J}$ with any $\mathbf{p}$, m, and ε, that is, it has a well-defined reduced form $\mathbf{x}_{-J} = h(\mathbf{p}, m, \varepsilon)$. Strict concavity and quasi-concavity for each ε imply the Weak Axiom, given ε. Together with strong monotonicity and continuity, this implies that the stochastic demand functions are homogeneous of degree zero and satisfy Walras' Law almost surely (a.s.) Homogeneity, Walras' Law, and the Weak Axiom, in turn, imply negative definiteness of the Slutsky matrix a.s., while almost sure Slutsky symmetry is implied by maximization of random utility.[3] The preservation of these properties is an important feature of the class of random utility models considered in this paper, because it implies that, in the terminology of Browning and Carro (2007), parameters and data are variation independent.

Assumption A3. *The $(J-1) \times (J-1)$ matrix $\nabla_\varepsilon \mathbf{MRS}(\mathbf{x}, \varepsilon)$ has full rank $J-1$ for all ε.*

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 $\mathbb{R}^J$.

3. For a detailed discussion see Mas-Colell, Winston and Green (1995).

Assumption A4. $\dim(\varepsilon) = J - 1$, *and* P_ε *is functionally independent of* $(\boldsymbol{p}', m)$.

Assumptions A1 and A2 imply that the Jacobian of h with respect to ε exists. Assumption A3 is necessary for it to have full rank a.s.; see Section 4. Assumption A4 is conventionally motivated by postulating that ε are utility parameters, and preferences are assumed not to depend on prices and income.[4] Together with Assumption A4, Assumptions A1–A3 guarantee that $\mathbf{x}_{-J}$ has a non-degenerate distribution on the budget set $B_{-J}(\mathbf{p}, m) = \{\mathbf{x}_{-J} \in \mathbb{R}_+^{J-1} : \mathbf{p}'\mathbf{x}_{-J} = m, x_J \geq 0\}$, given $\mathbf{p}$ and m. Conversely, since, given $\mathbf{p}$ and m, the co-domain of h is a $J-1$-dimensional space, the continuous invertibility of h with respect to ε requires that the dimension of ε be $J-1$. For h smooth with respect to ε, as a consequence of A2, Beckert (2006) shows that this dimensionality requirement is also *necessary* for the induced distribution of $\mathbf{x}_{-J}$ to be non-degenerate, given $\mathbf{p}$ and m.

On the basis of A1–A4, the probability space $(\mathbf{U}, \mathcal{U}, P_U)$ induces a probability space for demands. Let $(\mathbf{h}, \mathcal{H}, P_h)$ denote this probability space, where P_h is the probability measure induced by P_ε through the non-linear transformation $h(\mathbf{p}, m, \varepsilon)$, 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 $(\mathbf{U}, \mathcal{U}, P_U)$ and $(\mathbf{h}, \mathcal{H}, P_h)$ are consistent (or $\mathbf{h}$ is $\mathbf{U}$-rational), if i.a. for any $\mathbf{x}_{-J} \in \mathbb{R}_+^{J-1}$ satisfying $\mathbf{p}'\mathbf{x} \leq m$, $x_J \geq 0$, the inverse image of $\mathbf{x}_{-J} = h(\mathbf{p}, m, \varepsilon)$ with respect to ε, given $\mathbf{p}$ and m, is in $\mathcal{U}$, that is, $P_h(h(\mathbf{p}, m, \varepsilon)) = P_U(\tilde{U}(\mathbf{p}, m, \mathbf{x}_{-J}))$, where $\tilde{U}(\mathbf{p}, m, \mathbf{x}_{-J})) = \{U \in \mathbf{U} : (\mathbf{x}_{-J}, m - \mathbf{p}'\mathbf{x}_{-J})' = (h(\mathbf{p}, m, \varepsilon), x_J)' = \arg\max_{\mathbf{p}'\mathbf{x} \leq m} U(\mathbf{x}, \varepsilon)\} \in \mathcal{U}$. In order for unambiguous revelation of stochastic preferences from stochastic demands, this inverse should be unique. The invertibility conditions provided in this paper guarantee such uniqueness.

3. GLOBAL INVERTIBILITY WITH MRS-SEPARABLE HETEROGENEITY

This section examines the specification of structural models in which unobserved preference heterogeneity ε enters the marginal rates 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—MRS-separable heterogeneity. There are clear advantages of such specifications. 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. Moreover, it will be seen that they readily induce global invertibility. 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 assumption is maintained:

Assumption A5. $\mathbf{MRS}(\mathbf{x}, \varepsilon)$ *is multiplicatively separable with respect to* ε*:*

$$\mathbf{MRS}(\mathbf{x}, \varepsilon) = \upsilon(\mathbf{x}) + K(\mathbf{x})\varepsilon,$$

where $\upsilon(\mathbf{x})$ *is a* $(J-1) \times 1$ *vector of non-negative functions,* $K(\mathbf{x})$ *is a* $(J-1) \times (J-1)$ *matrix with full rank and span equal to* R_{++}^{J-1}.

Note that A5 implies the previous assumption A3.

Theorem 3.1. *Suppose A1, A2, A4, and A5 hold. Then, for any* $\mathbf{p}$ *and* m, $h(\mathbf{p}, m, \varepsilon)$ *is globally invertible for all demand vectors* $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$*, and, hence,* $\mathbf{x}_{-J}$ *has a non-degenerate distribution on* $B_{-J}(\mathbf{p}, m)$*, given any* $\mathbf{p}$ *and* m.

4. Prices and income are treated as non-stochastic in this analysis. When they are stochastic, this assumption is usually replaced by a conditional independence assumption; see Brown and Walker (1989), Lewbel (2001).

Proof. From the first-order conditions and A5, it follows immediately that

$$\varepsilon = K(\mathbf{x})^{-1}(\mathbf{p} - \upsilon(\mathbf{x})),$$

where $\mathbf{x} = (\mathbf{x}'_{-J}, x_J)'$, $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$, and $x_J = m - \mathbf{p}'\mathbf{x}_{-J}$ are well defined by A1 and A2. Non-degeneracy follows from A4. ||

Two examples can serve as illustrations of this result.

Example 3.1. *Consider the random utility model*

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

where $u(\cdot)$ is defined on $\mathbb{R}_+^{J-1}$, is monotonically increasing, and is weakly concave, $\nu(\cdot)$ is defined on $\mathbb{R}_+^J$ and satisfies A1, *and ε has a non-degenerate distribution on $\mathbb{R}_+^{J-1}$* (A4). *In this model, preferences are non-separable over the J goods, and marginal utilities may involve any subset of the components of ε. Concavity of $u(\cdot)$ and strict concavity of $\nu(\cdot)$ imply that* A2 *holds.*[5] *Then*

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

where $\upsilon(\mathbf{x}) = \left[\frac{\frac{\partial}{\partial x_j}\nu(\mathbf{x})}{\frac{\partial}{\partial x_J}\nu(\mathbf{x})}\right]_{j=1,\dots,J-1} \in \mathbb{R}_+^{J-1}$ *and* $K(\mathbf{x}) = \left[\frac{\partial}{\partial x_J}\nu(\mathbf{x})\right]^{-1}\left[\frac{\partial}{\partial x_j}u(\mathbf{x}_{-J})'\right]_{j=1,\dots,J-1}$. *A1 and A2 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 $u(\cdot)$ the identity function, that is, $u(\mathbf{x}_{-J}) = \mathbf{x}_{-J}$ for any $\mathbf{x}_{-J} \in \mathbb{R}_+^{J-1}$, and $\nu(\mathbf{x}) = \phi(\mathbf{x}) + x_J$, so that $U(\mathbf{x}, \varepsilon) = \phi(\mathbf{x}) + \mathbf{x}'_{-J}\varepsilon + x_J$.[6] *Brown and Matzkin's model implies that marginal rates of substitution are additive in ε, hence invertibility follows also directly from the first-order conditions. Note that this class of models, unlike Brown and Matzkin's, encompasses utility with heterogeneous curvature.*

Example 3.2. *This example combines CES features with Barten-type equivalence scaling.*[7] *For the case of three goods, with $\boldsymbol{x} = [x_1, x_2, x_3]'$ and $\varepsilon = [\varepsilon_1, \varepsilon_2]'$, suppose that*

$$U(\boldsymbol{x}; \varepsilon, \rho) = \left(\varepsilon_1 x_1^\rho + \varepsilon_2 x_2^\rho + x_3 + \left(\frac{x_1}{2} + \frac{x_2}{2}\right)^\rho\right)^{1/\rho},$$

where $\rho \in (0, 1)$ is a fixed parameter, and $\{\varepsilon_i > 0, i = 1, 2\}$ are random components. Then, $\mathbf{MRS}(\boldsymbol{x}, \varepsilon, \rho) = vec\{\rho\varepsilon_i x_i^{\rho-1} + \frac{\rho}{2}(\frac{x_1}{2} + \frac{x_2}{2})^{\rho-1}, i = 1, 2\}$, *while* $\nabla_{x_3}\mathbf{MRS}(\boldsymbol{x}, \varepsilon, \rho) = \boldsymbol{0}$, *and* $\nabla_\varepsilon \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \rho) = \mathrm{diag}\{\rho x_i^{\rho-1}, i = 1, 2\}$. *Hence,*

$$\upsilon(\boldsymbol{x}) = \iota\frac{\rho}{2}\left(\frac{x_1}{2} + \frac{x_2}{2}\right)^{\rho-1}, \quad \textit{where } \iota = [1, 1]',$$

$$K(\boldsymbol{x}) = \mathrm{diag}\{\rho x_i^{\rho-1}, i = 1, 2\},$$

5. To see this, note that $\nabla_{\mathbf{xx}'}U(\mathbf{x}, \varepsilon)$ is an M-matrix, as defined in Horn and Johnson (1991), $\nabla_{\mathbf{x}\varepsilon'}U(\mathbf{x}_{-J}, \varepsilon)$ is a diagonal matrix, $\nabla_{x_J\varepsilon'}U(\mathbf{x}, \varepsilon) = \mathbf{0}$, and $\nabla_{\varepsilon\varepsilon'}U(\mathbf{x}, \varepsilon) = \mathbf{0}$, while $\nabla_{\mathbf{w}}U(\mathbf{x}, \varepsilon) >> 0$ for $\mathbf{w} = (\mathbf{x}'\varepsilon')'$.

6. Note that these restrictions do not amount to a positive monotonic transformation, so that these two random utility models do not belong to the same equivalence class.

7. See Barten (1964, 1968).

so that, given $\boldsymbol{p} = [p_1, p_2]'$*, for* $\boldsymbol{x}_{-3} \in B_{-3}(\boldsymbol{p}, m)$*, inversion yields*

$$(\varepsilon_1, \varepsilon_2)' = diag\{x_i^{1-\rho}/\rho, i = 1, 2\}\left[\boldsymbol{p} - \iota\frac{\rho}{2}\left(\frac{x_1}{2} + \frac{x_2}{2}\right)^{\rho-1}\right].$$

As these examples show, there exists a class of random utility models, which are non-trivial generalizations of the linear heterogeneity model and are globally invertible. This class induces marginal rates of substitution with multiplicatively separable heterogeneity and heterogeneous utility curvature.

4. GLOBAL INVERTIBILITY WITHOUT MRS-SEPARABLE HETEROGENEITY

This section relaxes the MRS-separable heterogeneity restriction and treats the general non-separable case. The exposition illuminates the power of the restrictions that the multiplicative separability assumption of the preceding section imposes. It starts with a discussion of the distinct requirements for local and global homeomorphisms, respectively and it subsequently presents a sequence of results for general, non-separable cases that yield global homeomorphic demands, given $\mathbf{p}$ and m.

4.1. *Local vs. global homeomorphisms*

With multiplicatively separable heterogeneity in the marginal rates of substitution, global invertibility follows immediately from the first-order conditions. In general, however, the homeomorphism property must be deduced from properties of the Jacobian $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$, a $(J-1) \times (J-1)$ matrix. Since the first-order conditions of the utility maximization problem hold for every ε, this Jacobian can be expressed as

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

where $\mathbf{x} = (h(\mathbf{p}, m, \varepsilon)', m - h(\mathbf{p}, m, \varepsilon)'\mathbf{p})'$, for any $\mathbf{p}, m, \varepsilon$. Assumption A2, in its conventional form with respect to $\mathbf{x}$, implies that $\text{rk}(\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \varepsilon)) = J-1$ for $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m) = \{\mathbf{x}_{-J} \in \mathbb{R}_+^{J-1} : \mathbf{p}'\mathbf{x}_{-J} = m, x_J \geq 0\}$. Hence, $\text{rk}(\nabla_\varepsilon \mathbf{MRS}(x, \varepsilon)) = J-1$ is necessary for the Jacobian to have full rank. If it has full rank, then, together with h being continuous, this is sufficient for h to be a local homeomorphism with respect to ε, by an application of the Implicit Function Theorem. Note that the Implicit Function Theorem involves sufficient, not necessary conditions for local invertibility.

Global invertibility places stronger conditions on the Jacobian. Arguments establishing global homeomorphisms rest on applications of the theorems by Gale and Nikaido (1965) or Mas-Colell (1979). Akin to the weaker Implicit Function Theorem, these theorems provide also only sufficient conditions for the existence of global homeomorphisms. They essentially strengthen the requirement from the Jacobian having full rank to the Jacobian having some or all principal minors positive, that is, they also place conditions on the principal sub-matrices of $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$. Specifically, the Gale and Nikaido (1965) conditions are: the support of ε is a rectangle, and the Jacobian is a P matrix for every ε, that is, every principal minor has positive sign. The Mas-Colell (1979) conditions are weaker; if the support of ε is a rectangle, the Jacobian needs to be a P matrix only at its vertices, and, for ε in the interior of its support, it is only required that the Jacobian have a positive determinant. An example of a function $\psi(\varepsilon)$ satisfying Gale and Nikaido (1965) conditions is provided in the Appendix.

Cast in this framework, the research question of this paper can be re-stated: what conditions on $\mathbf{MRS}(\mathbf{x}, \varepsilon)$ induce the Gale and Nikaido (1965) or Mas-Colell (1979) conditions for

$\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$? In light of the foregoing discussion, and recognizing that utility functions form equivalence classes whose members have the same marginal rates of substitution, it is sensible to proceed with assumptions on $\mathbf{MRS}(\mathbf{x}, \varepsilon)$, rather than on $U(\mathbf{x}, \varepsilon)$, in the main results of this section. These results will be illustrated with examples of random utility functions $U(\mathbf{x}, \varepsilon)$ that facilitate the interpretation of such assumptions.

As this paraphrase of the main tools to establish local and global homeomorphism has alluded to, within the constraints of these invertibility theorems there is unfortunately no scope to determine necessary conditions for global homeomorphisms.

The conditions presented in this paper relate to interior solutions of the consumer's utility maximization problem. Allowing for corner solutions permits slightly weaker conditions. Corner solutions, say $x_j = h_j(\mathbf{p}, m, \varepsilon) = 0$ for some $j = 1, \ldots, J$, are characterized by $\frac{\partial}{\partial x_j} U(\mathbf{x}, \varepsilon) \leq \lambda p_j$, while the Kuhn–Tucker condition of the constrained optimization problem

$$\mathbf{x}'(\nabla_{\mathbf{x}} U(\mathbf{x}, \varepsilon) - \lambda(\mathbf{p}', p_J)') = 0$$

continues to hold at $\mathbf{x} = (h(\mathbf{p}, m, \varepsilon)', m - h(\mathbf{p}, m, \varepsilon)'\mathbf{p})'$, for any ε, where $\lambda > 0$ is the shadow value of m. Assuming that one good is consumed in positive amounts a.s., say $x_J > 0$, it follows that

$$\mathbf{x}'_{-J}(\mathbf{MRS}(\mathbf{x}, \varepsilon) - \mathbf{p}) = 0 \quad \text{a.s.}$$

This implies that

$$\mathbf{0} = A(\mathbf{p}, m, \varepsilon) + B(\mathbf{p}, m, \varepsilon) \quad \text{a.s.,} \tag{4.1}$$

where

$$\begin{aligned} A(\mathbf{p}, m, \varepsilon) &= \nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)' (\mathbf{MRS}(\mathbf{x}, \varepsilon) - \mathbf{p}) \\ B(\mathbf{p}, m, \varepsilon) &= [\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)' + \nabla_\varepsilon \mathbf{MRS}(\mathbf{x}, \varepsilon)'(\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \varepsilon)')^{-1}] \times \cdots \\ &\qquad \nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \varepsilon)' h(\mathbf{p}, m, \varepsilon), \end{aligned}$$

for $\mathbf{x} = (h(\mathbf{p}, m, \varepsilon)', m - \mathbf{p}'h(\mathbf{p}, m, \varepsilon))'$. Hence, for interior solutions, $\mathbf{MRS}(\mathbf{x}, \varepsilon) = \mathbf{p}$, so that (4.1) implies that $B(\mathbf{p}, m, \varepsilon) = \mathbf{0}$ a.s. Strict concavity, in turn, yields $\text{rk}(\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \varepsilon)) = J - 1$ a.s., so that the Jacobian satisfies $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)' = -\nabla_\varepsilon \mathbf{MRS}(\mathbf{x}, \varepsilon)'(\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \varepsilon)')^{-1}$, as above. For corner solutions, $\mathbf{MRS}(\mathbf{x}, \varepsilon) - \mathbf{p} \neq \mathbf{0}$ on a set of positive P_ε measure, so that, in order for (4.1) to hold, $\nabla_\varepsilon \mathbf{MRS}(\mathbf{x}, \varepsilon)$ and, hence, the Jacobian $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$ may be rank deficient with positive probability. Therefore, in the presence of corner solutions, the conditions presented in this paper need only hold with positive probability, rather than almost surely.

4.2. *Global invertibility in the general case*

This section considers a variety of structural models where marginal rates of substitution are not separable in heterogeneity. Beckert (2006) provides a related result which requires strong symmetry and definiteness assumptions on the product of the derivatives of $\mathbf{MRS}(\mathbf{x}, \varepsilon)$ with respect to its arguments. The conditions provided here cover a wider class of models and are readily interpretable in terms of microeconomic consumer choice theory.

The following additional assumptions are maintained:

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

This assumption is innocuous since, within the framework of random utility models laid out in Section 2, the distribution of ε is independent of $\mathbf{p}$ and m. It is necessary when applying the Gale and Nikaido (1965) and Mas-Colell (1979) Theorems.

Assumption A6. *$U(\mathbf{x}, \varepsilon)$ is strictly concave in $\mathbf{x}_{-J}$ and linear in the outside good x_J, and $\nabla_\varepsilon \mathbf{MRS}(\mathbf{x}, \varepsilon)$ is positive definite a.s.*

Assumption A6 implies A3 and encompasses a wide class of random utility models and merely requires the existence of an outside good for which utility is linear. Often, such outside options have natural interpretations in applications, as they amount to the difference between m and the expenditure on the inside goods, $\mathbf{p}'h(\mathbf{p}, m)$.

Theorem 4.1. *Suppose that A1, A2, A4′, and A6 hold. Then, for any $\mathbf{p}$ and m, $h(\mathbf{p}, m, \varepsilon)$ 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. A6 implies that $-\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x}, \varepsilon)$ is positive definite for all $\mathbf{x}$ and ε, and symmetric. Its inverse inherits these properties. Horn and Johnson (1985), theorem 7.6.3, then implies that its product with a positive definite matrix $\nabla_\varepsilon \mathbf{MRS}(\mathbf{x}, \varepsilon)$ is diagonalizable, that is, similar[8] to a diagonal matrix, whose eigenvalues are positive. Similarity means that there exists a non-singular transformation S of $\mathbf{x}_{-J} = h(\mathbf{p}, m, \varepsilon)$, possibly dependent on $\mathbf{p}, m, \varepsilon$, 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

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

where $S(\mathbf{p}, m, \varepsilon)$ is a non-singular matrix consisting of the $J-1$ eigenvectors of $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$ and $D(\mathbf{p}, m, \varepsilon)$ is a diagonal matrix with the positive eigenvalues of $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$ on its diagonal. This is necessary and sufficient for $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$ to be positive definite almost surely. 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 k-th principal minor of the Jacobian matrix $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$ is

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

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

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

where $E_k' \mathbf{y} \neq 0$ and the last inequality follows because $\nabla_\varepsilon h(\mathbf{p}, m, \varepsilon)$ 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 $P_k(\mathbf{p}, m, \varepsilon)$ such that

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

Therefore, the Gale and Nikaido (1965) conditions are satisfied. ||

8. An $n \times n$ matrix A is similar to an $n \times n$ matrix B if there exists a non-singular $n \times n$ matrix S such that $B = S^{-1} A S$. Similarity is an equivalence relation. See Horn and Johnson (1985), section 1.3, for further details.

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

Assumption A7. *(i) $\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x},\varepsilon)$ has negative diagonal and non-negative off-diagonal entries, a.s.; (ii) $(-1)^J \nabla_\varepsilon \mathbf{MRS}(\mathbf{x},\varepsilon)$ has positive diagonal and non-positive off-diagonal entries, and all its principal minors are positive, a.s.; and (iii) $(-1)^J \nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x},\varepsilon) - \nabla_\varepsilon \mathbf{MRS}(\mathbf{x},\varepsilon) \geq \mathbf{0}$, a.s.*

Matrices, having properties as in (i) and (ii), are sometimes referred to as M-matrices; see, for example, Horn and Johnson (1991). Assumption A7 is particularly attractive because it has an economic interpretation. It can be shown that the restrictions imposed by A7 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. This structure may, for instance, prevail in demand for Internet services, where users choose between (the intensity of usage of) various web applications and bandwidth: the applications are possibly substitutable, but each can be performed more efficiently at higher levels of bandwidth. One can think of numerous other applications which share the generic features of this example: usage of substitutable varieties of "software" and complementary "hardware".

Let $\mathbf{Z}$ be the class of square matrices whose off-diagonal elements are all non-positive, as in Fiedler and Pták's (1962) definition (4,1). And let $\mathbf{K}$ be those elements in $\mathbf{Z}$ which have all principal minors positive, as in Fiedler and Pták's (1962) definition (4,4). Theorem 5.4 below uses Fiedler and Pták's (1962) theorem (4,6): If $A \in \mathbf{K}$, $B \in \mathbf{Z}$ and $B - A \geq \mathbf{0}$, then, i.a., $B^{-1}A \in \mathbf{K}$.

Theorem 4.2. *Suppose that A1, A2, A4′, and A7 hold. Then, for any $\mathbf{p}$ and m, $h(\mathbf{p},m,\varepsilon)$ 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 A7(i), $-\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x},\varepsilon)$ has positive diagonal and non-positive off-diagonal entries. Hence it belongs to the class $\mathbf{Z}$; and by A7(ii), $(-1)^J \nabla_\varepsilon \mathbf{MRS}(\mathbf{x},\varepsilon)$ belongs to the class $\mathbf{K}$. Hence, using A10(iii), by Fiedler and Pták (1962), theorem (4,6),

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

that is, all its principal minors are positive, so that the Gale and Nikaido (1965) conditions are satisfied. ||

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

Assumption A8. *$\nabla_\varepsilon \mathbf{MRS}(\mathbf{x},\varepsilon)$ and $-\left[\nabla_{\mathbf{x}_{-J}} \mathbf{MRS}(\mathbf{x},\varepsilon)\right]^{-1}$ are strictly totally positive and bounded a.s.,*

this establishes the following

Theorem 4.3. *Suppose that A1, A2, A4′, and 9 hold. Then, for any* $\mathbf{p}$ *and* m, $h(\mathbf{p}, m, \varepsilon)$ *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.

The following example is a variant of Example 3.2 and illustrates both Theorem 4.1 and Theorem 4.3.

Example 4.1. *Consider the random utility model for three goods* $\boldsymbol{x} = (x_1, x_2, x_3)'$,

$$U(\boldsymbol{x}, \varepsilon, \theta) = \left(\frac{\theta_1}{\varepsilon_1}(x_1+1)^{\varepsilon_1} + \frac{\theta_2}{\varepsilon_2}(x_2+1)^{\varepsilon_2} + x_3 + \frac{\theta_3}{\varepsilon_1+\varepsilon_2}\left(\frac{x_1}{2}+\frac{x_2}{2}+1\right)^{\varepsilon_1+\varepsilon_2}\right)^{1/\varepsilon_1+\varepsilon_2},$$

where $\theta = (\theta_1, \theta_2, \theta_3)'$ *is a vector of fixed, positive parameters, and* ε_1 *and* ε_2 *are random utility parameters, with a continuous distribution on the unit interval and satisfying* $\varepsilon_1 + \varepsilon_2 < 1$.[9] *In this case, heterogeneity does not enter the marginal rates of substitution in a separable fashion:*

$$\mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta) = vec\{\theta_i(x_i+1)^{\varepsilon_1 - 1}, i = 1, 2\} + \iota\frac{\theta_3}{2}\left(\frac{x_1}{2}+\frac{x_2}{2}+1\right)^{\varepsilon_1+\varepsilon_2-1}.$$

In this example,

$$\nabla_\varepsilon \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta) = \operatorname{diag}\{\theta_i(x_i+1)^{\varepsilon_i-1}\ln(x_i+1), i = 1, 2\} + c\iota\iota'$$

$$\nabla_{\boldsymbol{x}_{-3}} \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta) = \operatorname{diag}\{\theta_i(\varepsilon_i - 1)(x_i+1)^{\varepsilon_i-2}, i = 1, 2\} + d\iota\iota',$$

where $c = \frac{\theta_3}{2}\left(\frac{x_1}{2}+\frac{x_2}{2}+1\right)^{\varepsilon_1+\varepsilon_2-1}\ln\left(\frac{x_1}{2}+\frac{x_2}{2}+1\right) > 0$, *and* $d = \frac{\theta_3(\varepsilon_1+\varepsilon_2-1)}{4}\left(\frac{x_1}{2}+\frac{x_2}{2}+1\right)^{\varepsilon_1+\varepsilon_2-2} < 0$ *a.s. It is easy to verify that* $|\nabla_\varepsilon \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)| > 0$ *a.s., so that the positive diagonal elements of* $\nabla_\varepsilon \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)$ *imply that this matrix is strictly positive definite. Similarly, the support restrictions on* ε_i, $i = 1, 2$, *imply that* $|\nabla_{\boldsymbol{x}_{-3}} \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)| > 0$ *a.s., and therefore, a.s.,* $|-\left[\nabla_{\boldsymbol{x}_{-3}} \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)\right]^{-1}| > 0$. *The diagonal elements of* $-\left[\nabla_{\boldsymbol{x}_{-3}} \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)\right]^{-1}$ *are proportional*[10] *to* $-\theta_i(\varepsilon_i - 1)(x_i+1)^{\varepsilon_i-2} - d$, $i = 1, 2$, *and, as a consequence of the support restrictions on* ε, *also positive. Therefore,* $-[\nabla_{\boldsymbol{x}_{-3}} \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)]^{-1}$ *is also strictly totally positive a.s. For finite income* m *and prices* $\boldsymbol{p}$ *bounded away from zero, all matrix components are finite on* $B_{-3}(\boldsymbol{p}, m)$. *Then, the Cauchy–Binet formula implies that* $-[\nabla_{\boldsymbol{x}_{-3}} \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)]^{-1}\nabla_\varepsilon \mathbf{MRS}(\boldsymbol{x}, \varepsilon, \theta)$ *is strictly totally positive a.s.; that is, in particular it is a* P *matrix and therefore has all principal minors positive a.s.*

5. CONCLUSIONS

This paper fills an important gap in the theoretical microeconometric analysis of consumer choice. It provides conditions on structural non-parametric 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 non-parametrically identifiable random utility models with multiplicatively separable heterogeneity beyond the classical model of Brown and Matzkin (1995) and discusses various extensions to completely non-separable

9. Strictly speaking, the support of the joint distribution of $[\varepsilon_1, \varepsilon_2]$ is, therefore, a triangular set. This can be turned into a rectangle by requiring that $\varepsilon_i \in \left(0, \frac{1}{2}\right)$, $i = 1, 2$.

10. Up to a positive scale factor.

cases. These results and the illustrative examples provided throughout the paper broaden the class of random utility models suitable for parametric and non-parametric microeconometric analysis.

Applied researchers face the typical trade-off between tightly, often parametrically specified models, as, for example, the foregoing parametric examples, which by design satisfy regularity conditions implied by microeconomic theory; and more general, non-parametric specifications which proceed by assuming that these conditions hold. This trade-off extends to the global invertibility conditions provided in this paper. In the case of parametric models, estimation can naturally be carried out in a method of moments framework, possibly assisted by simulation methods (Beckert, 2005).

Non-parametric demand analyses may generally employ convenient and flexible non-parametric specifications, assuming that global invertibility holds, under some of the conditions given above. Alternatively, they can start from relatively flexible representations of random utility or random marginal rates of substitution, for example, building on the multiplicatively separable model, as in A5. Suppose, for example, that $K(\mathbf{x})\varepsilon$ in A5 has a polynomial approximation, that is, $K(\mathbf{x})\varepsilon = K(\mathcal{P};\mathbf{x})\varepsilon$, where $K(\mathcal{P};\mathbf{x})$ is a matrix polynomial in the exponentiation operator $\mathcal{P}$,

$$K(\mathcal{P};\mathbf{x}) = \mathbf{I}_{J-1} - \kappa(\mathcal{P};\mathbf{x}) = \mathbf{I}_{J-1} - \sum_{k=1}^{p} \kappa_k(\mathbf{x})\mathcal{P}^k,$$

the exponentiation operator is defined by $\mathcal{P}^k\varepsilon = [\varepsilon_j^k]_{j=1,\dots,J-1}$, $k = 1,\dots,p$, the polynomial degree p may be finite or infinite, and the coefficient matrices $\{\kappa_k(\mathbf{x}), k = 1,\dots,p\}$ can be flexible functions of $\mathbf{x}$. Then, assuming that, for any $\mathbf{x}_{-J} \in B_{-J}(\mathbf{p}, m)$ and $x_J = m - \mathbf{p}'\mathbf{x}_{-J}$, the characteristic equation $|\mathbf{I}_{J-1} - \kappa(z;\mathbf{x})| = 0$ has all $J-1$ roots outside the unit circle, the inverse matrix polynomial $K(\mathcal{P};\mathbf{x})^{-1}$ exists, and its coefficient matrices can be related to matrices $\{\kappa_k(\mathbf{x}), k = 1\cdots, p\}$ by well-known formulae.[11] This model provides added flexibility in terms of modelling unobserved heterogeneity, while retaining the model's global invertibility property.

Acknowledgements. We are indebted to Arthur Lewbel, Rosa Matzkin and Dan McFadden for insightful comments and discussions. We also thank the editor and two referees for helpful comments. Blundell is grateful for the financial support of the ESRC Centre for the Micro-economic Analysis of Public Policy at the Institute for Fiscal Studies.

APPENDIX. EXAMPLE: FUNCTION $\psi(\cdot)$ SATISFYING GALE AND NIKAIDO (1965) CONDITIONS

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

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

where $\mathbf{A}_{j.}$, $j = 1,\dots,J-1$, is the j-th column of $\mathbf{A}$.

Suppose $\mathbf{A}$ is triangular, with positive diagonal elements. Then, $\nabla_\varepsilon \psi(\varepsilon)$ is triangular as well and, since $\psi(\varepsilon)$ has positive elements a.s., has positive diagonal elements, and the same is true for every principal submatrix of $\mathbf{A}$ and $\nabla_\varepsilon \psi(\varepsilon)$. Hence, $\nabla_\varepsilon \psi(\varepsilon)$ and its principal submatrices have determinants which equal their traces and hence are positive. Consequently, $\psi(\varepsilon)$ also satisfies the weaker Mas-Colell (1979) conditions.

11. Technically, this is closely related to the invertibility of vector autoregressive and moving average processes in multivariate time series analysis.

REFERENCES

BARTEN, A. P. (1964), "Consumer Demand Functions Under Conditions of Almost Additive Preferences", *Econometrica*, **32**, 1–38.

BARTEN, A. P. (1968), "Estimating Demand Equations", *Econometrica*, **36** (2), 213–251.

BECKERT, W. (2005), "Estimation of Stochastic Preferences: An Empirical Analysis of Demand for Internet Services", *Review of Economics and Statistics*, **87** (3), 495–502.

BECKERT, W. (2006), "Specification and Identification of Stochastic Demand Models", *Econometric Reviews*, **26** (6), 669–683.

BECKERT, W. and BLUNDELL, R. (2005), "Heterogeneity and the Nonparametric Analysis of Consumer Choice: Conditions for Invertibility" (Working Papers CWP09/05, cemmap).

BENKARD, C. L. and BERRY, S. (2006), "On the Nonparametric Identification of Nonlinear Simultaneous Equations Models: Comment on B. Brown (1983) and Roehrig (1988)", *Econometrica*, **74** (5), 1429–1441.

BROWN, B. W. (1983), "The Identification Problem in Systems Nonlinear in the Variables", *Econometrica*, **51** (1), 175–196.

BROWN, B. W. and WALKER, M. B. (1989), "The Random Utility Hypothesis and Inference in Demand Systems", *Econometrica*, **57** (4), 815–829.

BROWN, D. J. and MATZKIN, R. L. (1995), "Estimation of a Random Utility Model from Data on Consumer Demand" (Mimeo, Yale University and Northwestern University).

BROWNING, M. and CARRO, J. (2007), "Heterogeneity and Microeconometrics Modelling", in R. Blundell, W. Newey and T. Persson (eds.) *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. 3, ch. 3, *Econometric Society Monographs*, **43** (Cambridge: Cambridge University Press) 47–74.

FIEDLER, M. and PTÁK, V. (1962), "On Matrices with Non-Positive Off-Diagonal Elements and Positive Principal Minors", *Czechoslovak Mathematical Journal*, **12**, 382–400.

GALE, D. and NIKAIDO, H. (1965), "The Jacobian Matrix and Global Univalence of Mappings", *Mathemathische Annalen*, **159** (2), 81–93.

HECKMAN, J. (2001), "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture", *Journal of Political Economy*, **109**, 673–748.

HORN, R. A. and JOHNSON, C. R. (1985), *Matrix Analysis* (Cambridge, MA: MIT Press).

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

IMBENS, G. and NEWEY, W. K. (2002), "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity" (Mimeo, Department of Economics, MIT).

KARLIN, S. (1968), *Total Positivity* (Stanford: Standford University Press).

LEWBEL, A. (2001), "Demand Systems With and Without Errors", *American Economic Review*, **91** (3), 611–618.

LEWBEL, A. (2007), "Modelling Heterogeneity", in R. Blundell, W. Newey and T. Persson (eds.) *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. 3, ch. 5, *Econometric Society Monographs*, **43** (Cambridge: Cambridge University Press) 111–121.

MAS-COLELL, A. (1979), "Homeomorphisms of Compact, Convex Sets and the Jacobian Matrix", *SIAM Journal of Mathematical Analysis*, **10** (6), 1105–1109.

MAS-COLELL, A., WINSTON, M. D. and GREEN, J. R. (1995), *Microeconomic Theory* (New York: Oxford University Press).

MATZKIN, R. L. (2003), "Nonparametric Estimation of Nonadditive Random Functions", *Econometrica*, **71** (5), 1339–1376.

MATZKIN, R. L. (2005), "Nonparametric Simultaneous Equations" (Mimeo, Department of Economics, Northwestern University).

MATZKIN, R. L. (2007), "Heterogeneous Choice", in R. Blundell, W. Newey and T. Persson (eds.) *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. 3, ch. 4, *Econometric Society Monographs*, **43** (Cambridge: Cambridge University Press) 75–110.

MCFADDEN, D. L. (2004), "Revealed Stochastic Preference: A Synthesis" (Mimeo, UC Berkeley).

MCFADDEN, D. L. and RICHTER, K. (1971), "On the Extension of a Set Function on a Set of Events to a Probability on the Generated Boolean σ-algebra" (Working Paper, University of California Berkeley).

MCFADDEN, D. L. and RICHTER, K. (1990), "Stochastic Rationality and Revealed Stochastic Preference", in J. Chipman, D. McFadden, K. Richter (eds.) *Preference, Uncertainty, and Rationality*, (Boulder: Westview Press) 161–186.

ROEHRIG, C. S. (1988), "Conditions for Identification in Nonparametric and Parametric Models", *Econometrica*, **56** (2), 433–447.