Projektbereich A

Discussion Paper No. A–496

Metonomy and Cross Section Demand

by

Igor V. Evstigneev*

Werner Hildenbrand*

Michael Jerison*

November 1995

*Financial support by Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn is gratefully acknowledged.

METONYMY AND CROSS SECTION DEMAND

By I. V. Evstigneev, W. Hildenbrand, and M. Jerison

Cross section consumer expenditure data are frequently used to make conclusions about consumer demand behavior. Such conclusions, however, can only be justified under certain assumptions, which are often left unstated in the empirical demand literature. An assumption of this type, the *metonymy* hypothesis, was stated rigorously and then exploited by Härdle, Hildenbrand and Jerison when analyzing the monotonicity property of aggregate demand functions. The purpose of the present paper is to examine the metonymy hypothesis in more detail. We prove that the distribution of demand vectors derived from a not necessarily metonymic population is identical to the distribution derived from some metonymic one. This implies, in particular, that the metonymy hypothesis cannot be rejected or confirmed on the basis of data from a single cross section.[1]

Keywords: Cross section demand; Consumer expenditure data, Metonymy hypothesis, Law of Demand.

[1]I. V. Evstigneev, Central Economics and Mathematics Institute, Academy of Sciences of Russia, Krasikova 32, Moscow 117418, Russia. E-mail: evstign@cemi.msk.su . Corresponding author.

W. Hildenbrand, Department of Economics, University of Bonn, Lennéstrasse 37, D-53113, Bonn, Germany

M. Jerison, Department of Economics, State University of New York at Albany, Albany, New York 12222, USA

1. Introduction

Economists often use cross section consumer expenditure data to support or reject hypotheses about consumer demand behavior. For example, the hypothesis that the consumers in a particular population have homothetic preferences is considered to be rejected by cross section data satisfying "Engel's Law" that richer consumers spend smaller shares of their budgets on food. But to reject homotheticity based on cross section data one must make an additional assumption about the way the preferences or demand functions of richer and poorer consumers are related. A standard assumption is that richer consumers demand on average what poorer consumers would have demanded if their incomes were as high as the rich. Without some such assumption it is possible for cross section data from a population of homothetic consumers to satisfy Engel's Law. Poorer consumers could simply have a stronger preference for food and could be consuming more than the richer consumers would have if their incomes were equally low.

This possibility is illustrated in Figure 1. The points represent the consumers' demand vectors and the arrows represent the derivatives of the consumers' demands with respect to income. Homotheticity requires that the arrows lie on rays from the origin. Richer consumers in the diagram buy no more food (good 2) than poorer consumers, so the food budget shares of the rich are smaller than those of the poor. Yet the shape of a cross section Engel curve for food (a regression of food consumption on income) for this population tells us nothing about how the individual consumers' demands vary with income. Fig. 1

The standard assumption used to relate cross section data with demand behavior is that all the consumers, rich or poor, have a common demand function up to a random error that is uncorrelated with income. This assumption appears throughout the empirical demand literature, though it is often left unstated. It is quite stringent, as Figure 1 suggests. If it is satisfied, then the hypothesis that the common demand function is homothetic could be rejected on the basis of a large enough sample of demand vectors from a single cross section satisfying Engel's Law.

The hypothesis of a common demand function is restrictive, but it is not needed in order to draw behavioral inferences from a single cross section. Weaker assumptions suffice, for example, the *metonymy* hypothesis introduced by Hildenbrand and Jerison and used in Härdle at al. (1991) and Hildenbrand (1994).[2] Under metonymy, the distribution of demand vectors of the consumers with income level x can be used to represent the distribution of demand vectors that slightly poorer consumers would have if their incomes were raised to x. More precisely, under metonymy, if the poorer consumers' incomes increased to x and the means and covariances of their demand vectors were approximated by the means and covariances of the demands of the consumers with income x, then the approximation errors would be uncorrelated with income. (Note that these errors are not required to be small.)

In Härdle et al. (1991) and Hildenbrand (1994), the metonymy hypothesis is used in order to obtain evidence from cross section supporting the Law of Demand. That "law" states roughly that prices and mean demands move in opposite directions if

[2] In the common language, metonymy is a figure of speech in which the name of one thing is used to name another related thing, e.g. "crown" for "king."

consumers' incomes are fixed. If $D(p)$ is the market demand vector at the price vector p, then the Law of Demand states that for arbitrary price vectors p and q,

$$(D(p) - D(q))(p - q) < 0, \tag{1}$$

so the mean demand function is a strictly decreasing monotone operator (Wald (1936), Hildenbrand (1983); for a mathematical theory of monotone operators see, e.g., Vainberg (1973)). The Law of Demand is often assumed in economic applications. It implies that the equilibrium in a distribution economy is unique and globally stable. Under metonymy, if the cross section consumer demand data exhibit a strong form of heteroskedasticity, then the consumers' demand vectors spread out as their incomes rise. This latter condition implies the Law of Demand. The strong form of heteroskedasticity has empirical support, see Härdle et al. (1991) and Hildenbrand (1994), but that support comes from cross section data. The metonymy hypothesis is needed in order for that data to have behavioral implications. For this reason it is important to consider the empirical content of the metonymy hypothesis. What type of data would be needed to support or reject it? The present paper shows that the metonymy hypothesis cannot be rejected on the basis of data from a single cross section. The result obtained proves a conjecture by W. Hildenbrand.

Before discussing the result, we need to describe the model. We will consider the demand behavior of a population of consumers. Each consumer is characterized by a demand function and income, and the population is represented by a joint distribution on the space of demand functions and incomes. Since we will be concerned with a single cross section, prices are held fixed. The joint distribution on demand functions and incomes determines a distribution of the consumers' demand vectors. We will ask what this induced distribution tells us about the original joint distribution of demand functions and incomes. The demand *functions* contain the behavioral information referred to above. A finite set of cross section demand data can be interpreted as a random sample drawn from the induced distribution on the consumers' demand vectors. So we can also ask what such a finite sample can tell us about the distribution of consumers' demand functions and incomes.

This paper shows that data from a single cross section cannot refute the metonymy hypothesis. This might not seem surprising. Any finite set of data in which the consumers all have different incomes can be generated as a sample from a metonymic population in which the consumers all have the same demand function. However the main result presented below is deeper. It shows that the entire induced distribution of consumer demand vectors is identical to the distribution derived from a metonymic population. Thus cross sections derived from metonymic populations are indistinguishable from cross sections that are not, no matter how large the sample.

This conclusion might seem in conflict with Härdle et al. (1991), which reports a test of the metonymy hypothesis. However that test makes use of additional observed consumer characteristics. Suppose that there are observed characteristics (e.g. age or sex) that are correlated with consumer demands given consumer incomes. Then correlation between these characteristics and consumer incomes in a single cross section could allow one to reject the hypothesis that the consumers' demand functions and incomes are independently distributed. One can also partition the population into subgroups with identical observed characteristics and then test the joint hypothesis that the entire population *and all of its subgroups* satisfy metonymy. The test can be performed by comparing certain statistics of the distribution of demand vectors for the whole population with the mean of those statistics for the subgroups, cf. Härdle et al.(1991).

Such a test simply raises the question how one can tell that a subgroup with identical observed characteristics is metonymic. The answer given below is that it is necessary to have more than a single cross section. This does not mean that the metonymy hypothesis is redundant. It is needed in order to use cross section data to estimate effects of changes in income . But the information that would allow one to test the metonymy hypothesis (e.g., a sequence of cross sections with fixed prices and varying income distributions) would also allow one to estimate the income effects directly.

The main result and its versions are stated in the next section. The proofs are in Sections 3 and 4. The idea of our argumentation resembles the known method of changing the probability space in the theory of stochastic processes, cf. Skorohod (1965). The proof of a central result consists, rougly speaking, in finding an appropriate parametrization of the space of agents' characteristics. To find such a parametrization, one has to solve a system of integral equations in Hilbert space. This system is solved by using some general methods of functional analysis.

2. The model and the main results

Let $X = [x_0, x_1]$ be a segment in $(0, \infty)$ and let V be the set of all continuously differentiable vector functions $v(x) = (v_1(x), \dots, v_n(x))$ of $x \in X$ with values in a Euclidean space R^n $(n \geq 2)$ such that

$$v_i(x) \geq 0, \ i \in \{1, \dots, n\}, \ x \in X, \tag{2}$$

and

$$\sum_{i=1}^{n} v_i(x) = x, \ x \in X. \tag{3}$$

For a measurable space $(B, \mathcal{B})$, we denote by $W(B \times X)$ the class of all vector functions $w : B \times X \to R^n$ such that $w(b, \cdot) \in V$ for each $b \in B$ and $w(\cdot, x)$ is $\mathcal{B}$-measurable for any $x \in X$ (thus $w(b, x)$ is jointly measurable in (b, x)).

In the model under consideration, we are given: a probability measure $\rho(dx)$ on the σ-algebra $\mathcal{B}_X$ of Borel subsets of X, a measurable space $(B, \mathcal{B})$, a function $g(b, x) = (g_1(b, x), \dots, g_n(b, x))$ in the class $W(B \times X)$, and a stochastic kernel $\lambda(x, \Delta)$ $(x \in X, \Delta \in \mathcal{B})$ [3]. Numbers in $X = [x_0, x_1]$ represent possible values of consumers' incomes. The measure $\rho\,(dx)$ specifies the income distribution in the given population of consumers. B is the space of agents' characteristics. The measure $\lambda\,(x, db)$ on the space B describes the distribution of consumers' characteristics in the subgroup with income level x. The vector $g\,(b, x) = (g_1\,(b, x)\,, ..., g_n\,(b, x))$ represents demand for n commodities of an individual whose consumer's characteristics are specified by the element b of the space B and whose income is x. Prices are supposed to be fixed, the price of each commodity being equal to 1 (the latter normalization condition does not lead to a loss in generality).

[3] A real-valued function $\lambda(x, \Delta)$ of $x \in X$ and $\Delta \in \mathcal{B}$ is called a stochastic kernel if $\lambda(x, \cdot)$ is a probability measure on $\mathcal{B}$ for each $x \in X$ and $\lambda(\cdot, \Delta)$ is a Borel measurable function for each $\Delta \in \mathcal{B}$.

We assume:

(A.1) For any $i, j \in \{1, \dots, n\}$, the function

$$x \mapsto \int_B \lambda(x, db) g_i(b, x) g_j(b, x), \tag{4}$$

is continuously differentiable in $x \in X$. (The integral in (4) is well-defined and finite for each $x \in X$.)

(A.2) For any $i \in \{1, \dots, n\}$, $x \in X$ and $b \in B$, we have

$$g_i(b, x) > 0. \tag{5}$$

If $(C, \mathcal{C})$ and $(C', \mathcal{C}')$ are measurable spaces, δ a measure on $\mathcal{C}$, and $h : (C, \mathcal{C}) \to (C', \mathcal{C}')$ a measurable mapping, then we denote by $h \circ \delta$ the image of the measure δ under the mapping h, i.e., the measure δ' on on $\mathcal{C}'$ satisfying

$$\int_{C'} \delta'(dc') z(c') = \int_C \delta(dc) z(h(c)) \tag{6}$$

for any bounded measurable function $z(c')$, $c' \in C'$. If equation (6) holds for all bounded measurable functions $z(\cdot)$, then this equation holds for all functions $z(\cdot)$ for which the integrals in (6) are well-defined.

A central result is as follows.

Theorem 1. *There exist a measurable space $(A, \mathcal{A})$, a function*

$$f(a, x) = (f_1(a, x), \dots, f_n(a, x))$$

in the class $W(A \times X)$ and a stochastic kernel $\mu(x, da)$ on $X \times \mathcal{A}$ such that

$$f(\cdot, x) \circ \mu(x, \cdot) = g(\cdot, x) \circ \lambda(x, \cdot), \quad x \in X, \tag{7}$$

and

$$\int_X \rho(dx) \frac{d}{dx} \int_B \mu(x, db) f_i(b, x) f_j(b, x) = \int_X \rho(dx) \int_A \mu(x, da) \frac{\partial}{\partial x} [f_i(a, x) f_j(a, x)] \qquad (8)$$

for all $i, j \in \{1, 2, ..., n\}$.

Theorem 1 states that the original model can be transformed to another one which yields the same distribution of demand vectors (for each subgroup of consumers with income level $x \in X$) and satisfies additional condition (8). This condition is called the *metonymy* hypothesis. Mathematically, hypothesis (8) means a possibility of interchanging (at least in the average with respect to the measure ρ) the operators of differentiation in x and integration with respect to the stochastic kernel $\mu(x, da)$. The operators in question are applied to the finite family of functions $f_i(a, x) f_j(a, x)$, where i and j range over $\{1, 2, ..., n\}$. When analyzing consumer demand behavior, one often needs to know the values of the derivatives in the right-hand side of (8). These derivatives cannot directly be estimated, however. The metonymy property enables one to establish relations between the unobservable variables in the right-hand side of (8) and the variables in the left-hand side of (8), which can be estimated on the basis of observable cross-section data.

Remark 1. Writing formula (8), we presume that all the integrals under consideration are well-defined and finite. In particular, the functions $\frac{\partial}{\partial x}[f_i f_j]$ involved in the formulation of the metonymy hypothesis are supposed to be integrable. Actually, in the course of the proof of Theorem 1, we construct a mapping $f = (f_1, ..., f_n)$ for which the functions $\frac{\partial}{\partial x}[f_i f_j]$ are not only integrable, but uniformly bounded.

Remark 2. From (7), we obtain

$$\int_A \mu(x,da) f_i(a,x) f_j(a,x) = \int_B \lambda(x,db) g_i(b,x) g_j(b,x), \tag{9}$$

and so equation (8) is equivalent to the equation

$$\int_X \rho(dx) \frac{d}{dx} \int_B \lambda(x,db) g_i(b,x) g_j(b,x) = \int_X \rho(dx) \int_A \mu(x,da) \frac{\partial}{\partial x}[f_i(a,x) f_j(a,x)], \tag{10}$$

provided (7) holds.

In Theorem 1, the space A and the mapping f may depend on all the data of the original model. One can get a somewhat stronger result, in which these space and mapping are universal. To formulate the result, consider the function space V defined at the beginning of this section. Let us introduce a metric, r_V , in V:

$$r_V\left(u\left(\cdot\right), v\left(\cdot\right)\right) = |u(x_0) - v(x_0)| + \sup_{x \in X} |\frac{d}{dx} u(x) - \frac{d}{dx} v(x)|, \quad u(\cdot), \;\; v(\cdot) \in V,$$

where $|\cdot|$ is the Euclidean norm in R^n. The space V is complete and separable with respect to this metric. Denote by $\mathcal{V}$ the Borel σ-algebra on V and define

$$F_i(v,x) = v_i(x), \;\; x \in X, \;\; v(\cdot) \in V, \;\; i \in \{1, \ldots, n\}. \tag{11}$$

The function $F(v,x) = (F_1(v,x), \ldots, F_n(v,x))$, regarded as a function of x, coincides with $v(x)$ and thus belongs to the class V. Furthermore, for each $x \in X$, the function $F(\cdot, x)$ is $\mathcal{V}$-measurable. Consequently, $F(\cdot,\cdot) \in W(V \times X)$.

Theorem 2. *There exists a stochastic kernel $\gamma(x,dv)$ on $X \times \mathcal{V}$ such that, for every $x \in X$,*

$$F(\cdot,x) \circ \gamma(x,\cdot) = g(\cdot,x) \circ \lambda(x,\cdot) \tag{12}$$

and

$$\int\limits_X \rho(dx)\frac{d}{dx}\int\limits_V \gamma(x,dv)F_i(v,x)F_j(v,x) = \int\limits_X \rho(dx)\int\limits_V \gamma(x,dv)\frac{\partial}{\partial x}[F_i(v,x)F_j(v,x)] \quad (13)$$

for all $i,j \in \{1,\ldots,n\}$.

Theorem 2, which is formally stronger than Theorem 1, is, in fact, an easy consequence of it. To show this, consider the mapping $\pi : A \to V$ which associates to each $a \in A$ the function $f(a,\cdot) \in V$, constructed in Theorem 1. The mapping $\pi : (A,\mathcal{A}) \to (V,\mathcal{V})$ is measurable because, for every $x \in X$,

$$\frac{d}{dx}f(a,x) = \lim_{n\to\infty} n[f(a,x+\frac{1}{n}) - f(a,x)]$$

is an $\mathcal{A}$-measurable function. Consider a stochastic kernel $\gamma(x,dv)$ on $X \times \mathcal{V}$ defined by

$$\gamma(x,\cdot) = \pi \circ \mu(x,\cdot).$$

For every bounded Borel function $z(r)$, $r \in R^n$ we have

$$\begin{aligned}&\int\limits_V \gamma(x,dv)z(F(v,x)) = \int\limits_A \mu(x,da)z(F(\pi(a),x)) = \\ &\int\limits_A \mu(x,da)z(f(a,x)) = \int\limits_B \lambda(x,db)z(g(b,x))\end{aligned}$$

by virtue of (7) and the definitions of γ, F, and π. The last formula implies (12). In order to obtain (13), it remains to observe that

$$\int\limits_V \gamma(x,dv)\frac{\partial}{\partial x}[F_i(v,x)F_j(v,x)] = \int\limits_A \mu(x,da)\frac{\partial}{\partial x}[F_i(\pi(a),x)F_j(\pi(a),x)] =$$

$$\int\limits_A \mu(x,da)\frac{\partial}{\partial x}[f_i(a,x)f_j(a,x)]$$

by virtue of the definitions of γ, π, and F. Thus, we have shown that Theorem 2 is a consequence of Theorem 1.

To prove Theorem 2, we define

$$A = B \times X, \ \mathcal{A} = \mathcal{B} \times \mathcal{B}_X \tag{14}$$

and

$$\mu(x, \cdot) = \alpha_x \circ \lambda(x, \cdot),$$

where $\alpha_x : B \to A$ is the mapping transforming an element b of the space B into the element $\alpha_x(b) = (b, x)$ of the space A. For any $x \in X$ and any bounded measurable function $z\,(a)\,,\ a \in A$, we have

$$\int\limits_A \mu\,(x, da)\, z\,(a) = \int\limits_B \lambda\,(x, db)\, z\,(\alpha_x\,(b)) = \int\limits_B \lambda\,(x, db)\, z\,(b, x)\,. \tag{15}$$

Thus, the measure $\mu(x, da)$ is concentrated on the set

$$A_x = \{(b, y) \in A : y = x\},$$

and on this set (whose points can be identified with points of B), this measure has the same structure as the measure $\lambda(x, db)$ on B.

Theorem 2 is a consequence of the following

Theorem 3. *For the space $(A, \mathcal{A})$ and the stochastic kernel $\mu(x, da)$ defined by (14) and (15), there exists a vector function $f(a, x) = (f_1(a, x), \ldots, f_n(a, x))$ in $W(A \times X)$ satisfying (7) and (8).*

A proof of this statement will be given in Section 4.

3. Some systems of equations in L_2

This section contains some auxiliary results which are used in Section 4 to prove Theorem 3. Let us consider the Hilbert space $L_2(B \times X)$ of real-valued functions $\psi(b,x)$ of $b \in B$, $x \in X$ which are square-integrable with respect to the measure

$$m(db,dx) = \rho(dx)\lambda(x,db).$$

For ψ, $\psi' \in L_2(B \times X)$, we write

$$\langle \psi, \psi' \rangle = \int m(db,dx)\psi(b,x)\psi'(b,x).$$

Observe that every function $g_i(b,x)$, $i \in \{1,...,n\}$, is bounded, because $0 \le g_i(b,x) \le x \le x_1$ (where x_1 is the right end of the segment $X = [x_0, x_1]$). Consequently, $g_i \in L_2(B \times X)$. Define

$$\Phi(g_i, g_j) = \int\limits_X \rho(dx) \frac{d}{dx} \int\limits_B \lambda(x,db) g_i(b,x) g_j(b,x).$$

Proposition 1. *There exist bounded functions $\psi_1, \ldots, \psi_n \in L_2(B \times X)$ such that $\psi_i(b,x)$ is continuous in $x \in X$ for all $i \in \{1,\ldots,n\}$ and $b \in B$;*

$$\sum_{i=1}^{n} \psi_i(b,x) = 1, \; b \in B, \; x \in X; \tag{16}$$

and

$$\Phi(g_i, g_j) = \langle g_i, \psi_j \rangle + \langle g_j, \psi_i \rangle \tag{17}$$

for all, $i, j \in \{1,\ldots,n\}$.

The functions ψ_i, satisfying (16) and (17), will be used in the proof of Theorem 3 (see Section 4). They will represent the derivatives $\frac{\partial}{\partial x} f_i(a,x)$ of the functions f_i described in Theorem 3.

The proof of Proposition 1 relies on Lemmas 1 and 2 below.

For every l, $1 \le l \le n$, we consider an auxiliary system of equations

$$(\mathcal{S}_l) \quad \begin{cases} \Phi(g_i, g_j) = \langle g_i, \psi_j \rangle + \langle g_j, \psi_i \rangle, \ i,j \in \{1,\ldots,l\}, & (18) \\ \\ \Phi(g_j) = \langle xe, \psi_j \rangle, \ j \in \{1,\ldots,l\}, & (19) \end{cases}$$

where e is the function on $B \times X$ which is identically equal to 1 and

$$\Phi(g_j) = \int\limits_X \rho(dx) x \frac{d}{dx} \int\limits_B \lambda(x, db) g_j(b,x). \tag{20}$$

Note that the expression (20) is well-defined, since the function

$$\int\limits_B \lambda(x,b) g_j(b,x) = \frac{1}{x} \sum_{i=1}^{n} \int \lambda(x,b) g_i(b,x) g_j(b,x) \tag{21}$$

is continuously differentiable in $x \in X$ by virtue of (A.1).

Lemma 1. *Suppose that* $\{\psi_1, \ldots, \psi_l\} \subseteq L_2(B \times X)$ $(1 \le l \le n-1)$ *is a solution to system* $(\mathcal{S}_l)$ *and there exist real numbers* $r, r_1, \ldots, r_l$ *such that*

$$g_{l+1} = rxe + r_1 g_1 + \ldots + r_l g_l \tag{22}$$

almost everywhere on $X \times B$ *with respect to the measure* $m(db, dx)$*. Then the set of functions*

$$\psi_1, \ldots, \psi_l, \ \psi_{l+1} = re + \sum_{k=1}^{l} r_k \psi_k \tag{23}$$

is a solution to system $(\mathcal{S}_{l+1})$.

Proof. We write

$$\langle xe, \psi_{l+1}\rangle = \langle xe, re\rangle + \sum_{k=1}^{l} r_k \langle xe, \psi_k e\rangle =$$

$$\langle xe, re\rangle + \sum_{k=1}^{l} r_k \Phi(g_k) = \Phi(rxe) + \sum_{k=1}^{l} r_k \Phi(g_k) = \Phi(g_{l+1}),$$

because

$$\Phi(rxe) = \int_X \rho(dx) x \frac{d}{dx} \int_B \lambda(x, db) rx = \int_X \rho(dx) rx = \langle xe, re\rangle.$$

This proves (19) for $j = l+1$.

For $j \in \{1, \dots, l\}$, we have

$$\Phi(g_{l+1}, g_j) = \Phi(rxe, g_j) + \sum_{k=1}^{l} r_k \Phi(g_k, g_j) =$$

$$r\Phi(g_j) + \langle g_j, re\rangle + \sum_{k=1}^{l} r_k \langle \psi_k, g_j\rangle + \sum_{k=1}^{l} r_k \langle g_k, \psi_j\rangle =$$

$$r\Phi(g_j) + \langle \psi_{l+1}, g_j\rangle + \langle g_{l+1} - rxe, \psi_j\rangle = \langle \psi_{l+1}, g_j\rangle + \langle g_{l+1}, \psi_j\rangle,$$

since

$$\Phi(rxe, g_j) = \textstyle\int_X \rho(dx) \frac{d}{dx} \int_B \lambda(x, db) rx g_j(b, x) =$$

$$\int_X \rho(dx) rx \frac{d}{dx} \int_B \lambda(x, db) g_j(b, x) +$$

$$\textstyle\int_X \rho(dx) r \int_B \lambda(x, db) g_j(b, x) = r\Phi(g_j) + \langle g_j, re\rangle, \tag{24}$$

and $r\Phi(g_j) = r\langle xe, \psi_j\rangle$ by virtue of (19).

It remains to prove that

$$\Phi(g_{l+1}, g_{l+1}) = 2\langle g_{l+1}, \psi_{l+1}\rangle. \tag{25}$$

The left-hand side of (25) equals

$$\Phi(rxe,rxe)+2\sum_{k=1}^{l}r_k\Phi(rxe,g_k)+\sum_{k,i=1}^{l}r_kr_i\Phi(g_k,g_i)=$$

$$2r^2\langle xe,e\rangle+2\sum_{k=1}^{l}r_k\langle rxe,\psi_k\rangle+2\sum_{k=1}^{l}r_k\langle g_k,re\rangle+$$

$$\sum_{k,i=1}^{l}[r_kr_i\langle g_k,\psi_i\rangle+r_kr_i\langle g_i,\psi_k\rangle], \tag{26}$$

which follows from formulas (24), (18), (19), and from the identity

$$\Phi(rxe,rxe)=\int_X\rho(dx)\frac{d}{dx}\int_B\lambda(x,db)r^2x^2=2r^2\int_X\rho(dx)x.$$

On the other hand,

$$2\langle g_{l+1},\psi_{l+1}\rangle=2\langle rxe,re\rangle+2\sum_{k=1}^{l}\langle rxe,r_k\psi_k\rangle+$$

$$2\sum_{k=1}^{l}\langle r_kg_k,re\rangle+2\sum_{k,i=1}^{l}r_kr_i\langle g_k,\psi_i\rangle.$$

This, combined with (26), yields (25).

The lemma is proved.

Lemma 2. *Let $l\in\{1,...,n-1\}$. If the functions*

$$h_1=g_1,\dots,h_l=g_l,\;\; h_{l+1}=xe \tag{27}$$

are linearly independent, then system $(\mathcal{S}_l)$ has a solution $\{\psi_1,\dots,\psi_l\}$ such that the functions ψ_i are linear combinations of $h_1,...,h_{l+1}$ (and so ψ_i are bounded).

Proof. Suppose we have found a collection of functions $\{\psi_1,\dots,\psi_l\}$ which is a solution to the system

$$(\mathcal{S}'_l) \qquad \begin{cases} \frac{1}{2}\Phi(g_i, g_j) = \langle h_i, \psi_j \rangle, \; i, j \in \{1, \ldots, l\}, \\ \\ \Phi(g_j) = \langle h_{l+1}, \psi_j \rangle, \; j \in \{1, \ldots, l\}, \end{cases}$$

Then $\{\psi_1, \ldots, \psi_l\}$ is a solution to $(\mathcal{S}_l)$. In order to solve $(\mathcal{S}'_l)$, we fix $j \in \{1, \ldots, l\}$ and construct a function $\psi = \psi_j$ which represents a linear combination of $\{h_i\}$

$$\psi = \kappa_1 h_1 + \ldots + \kappa_{l+1} h_{l+1}$$

and satisfies

$$\frac{1}{2}\Phi(g_i, g_j) = \langle h_i, \psi \rangle, \; \Phi(g_j) = \langle h_{l+1}, \psi \rangle$$

$(i \in \{1, \ldots, l\})$. Such a function exists, because the Gram matrix $(\langle h_i, h_j \rangle)$, $i, j \in \{1, \ldots, l+1\}$, of the linearly independent family of functions (27) is non-degenerate.

The proof is complete.

Proof of Proposition 1. Suppose that the rank of the family of functions

$$xe, g_1, \ldots, g_{n-1} \tag{28}$$

is equal to l, where $1 < l \leq n$. Without loss of generality, we can assume that the functions $xe, g_1, \ldots, g_l$ are linearly independent and $g_{l+1}, \ldots, g_{n-1}$ can be expressed as linear combinations of $xe, g_1, \ldots, g_l$ (otherwise, we can renumerate the functions g_i). By virtue of Lemma 2, the system $(\mathcal{S}_l)$ has a solution, $\{\psi_1, \ldots, \psi_l\}$. Using Lemma 1 and representing, step by step, the functions $g_{l+1}, \ldots, g_{n-1}$ as linear combinations of $xe, g_1, \ldots, g_l$, we extend $\{\psi_1, \ldots, \psi_l\}$ to a solution $\{\psi_1, \ldots, \psi_{n-1}\}$ of the system $(\mathcal{S}_{n-1})$. We now observe that $g_n(x) = x - g_1(x) - \ldots - g_{n-1}(x)$, and, applying Lemma

1 once again, we obtain that there exist a solution $\{\psi_1, \ldots, \psi_n\}$ to the system $(\mathcal{S}_n)$ which satisfies

$$\psi_n = 1 - \psi_1 - \ldots - \psi_{n-1}.$$

This yields (16), and (18) with $l = n$ implies (17). The functions ψ_i we have constructed are linear combinations of xe and g_i (see Lemmas 1 and 2), and so ψ_i are bounded.

Thus, Proposition 1 is proved in the case when the rank of the system of functions (28) is greater than 1. If this rank equals 1, then, for some constants $r_1, r_2, \ldots, r_n$, we have

$$r_1 + \ldots + r_n = 1,$$

$$g_1(b,x) = r_1 xe, \ldots, g_n(b,x) = r_n xe,$$

for m-almost all (b,x). In this special case,

$$\Phi(g_i, g_j) = \int_X \rho(dx) \frac{d}{dx} \int_B \lambda(x, db) r_i r_j x^2 = 2 r_i r_j \int_X \rho(dx) x,$$

and equations (17) have the form

$$2 r_i r_j \int_X \rho(dx) x = \langle xe, r_i \psi_j + r_j \psi_i \rangle$$

$(i, j \in \{1, \ldots n\})$. These equations, as well as (16), are satisfied for $\psi_i = r_i e$, $i = 1, \ldots, n$.

Proposition 1 is proved.

4. A metonymic representation

To complete the proof of Theorem 3, we need the following proposition.

Proposition 2. *For all* $r \in X$, $q = (q_1, \ldots, q_n) \in R^n$, $p = (p_1, \ldots, p_n) \in R^n$ *satisfying*

$$q_i > 0, \ i \in \{1, \ldots, n\}, \tag{29}$$

$$\sum_{i=1}^{n} q_i = r, \ \sum_{i=1}^{n} p_i = 1, \tag{30}$$

there exists a function $\Xi(r, q, p; x)$ *of* $x \in X$ *such that* $\Xi(r, q, p; \cdot) \in V$, *the equations*

$$\Xi(r, q, p; r) = q, \tag{31}$$

$$\frac{\partial}{\partial x} \Xi(r, q, p; x)|_{x=r} = p \tag{32}$$

hold and, for each $x \in X$, *the function* $\Xi(r, q, p; x)$ *is Borel measurable with respect to* (r, q, p).

The function $\Xi(r, q, p; \cdot) \in V$ described above depends measurably on the parameters r, q, p and has the given value q and the given derivative p at the point $x = r$.

Proof of Proposition 2. Let us construct a function Ξ with the desired properties. Define the ith ($i = 1, \ldots, n-1$) coordinate of the vector $\Xi = (\Xi_1, \ldots, \Xi_n)$ by the formula

$$\Xi_i(r, q, p; x) = x[t \sin(s_i(x - r)) + \frac{q_i}{r}], \tag{33}$$

where

$$t = \frac{1}{2(n-1)r} \min\{q_1, \ldots, q_n\}, \ s_i = \frac{p_i - r^{-1} q_i}{tr}. \tag{34}$$

The nth component of the vector Ξ is defined by

$$\Xi_n = x - \Xi_1 - \ldots - \Xi_{n-1}. \tag{35}$$

Relations (31), (32) are easily obtained from (33)–(35) and (30). The function Ξ is smooth in x and Borel in (r,q,p). Thus, it remains only to show that $\Xi_i > 0$ for each $i = 1, \dots, n$. If $i < n$, then the positivity of Ξ_i follows from the fact that $t < \frac{q_i}{r}$. To prove the inequality $\Xi_n > 0$, we write

$$\Xi_n = x - \sum_{i=1}^{n-1} \Xi_{n-1} = x[(1 - \sum_{i=1}^{n-1} \frac{q_i}{r}) - t \sum_{i=1}^{n-1} \sin(s_i(x-r))] \geq x[\frac{q_n}{r} - t(n-1)] > 0.$$

The proof is complete.

Let us deduce Theorem 3 from Propositions 1 and 2.

Proof of Theorem 3. For $a = (b,y) \in A$ we define the function $f(a,x) = (f_1(a,x), ..., f_n(a,x))$ by the formula

$$f(a,x) = \Xi(y, g(b,y), \psi(b,y); x), \tag{36}$$

where Ξ is the function constructed in Proposition 2 and $\psi = (\psi_1, \dots, \psi_n)$ is the collection of functions described in Proposition 1. Observe that $f(a,x)$ is well-defined, since $\sum \psi_i = 1$ and the vector $q = g(b,y)$ satisfies conditions $\sum g_i(b,y) = y$, $g_i(b,y) > 0$ by virtue of $(A.2)$. (This is the only point in the proof, where condition $(A.2)$ is used.) For every $a \in A$, the function $f(a,x)$ belongs to V, and for every $x \in X$ it is measurable with respect to $a = (b,y)$ because $g(b,y)$ and $\psi(b,y)$ are measurable with respect to (b,y) and $\Xi(r,q,p)$ is measurable with respect to (r,q,p). Thus $f(\cdot,\cdot) \in W(A \times X)$.

If $a = (b,x)$, then $f(a,x) = \Xi(x, g(b,x), \psi(b,x); x)$, and so

$$f_i(a,x) = g_i(b,x), \tag{37}$$

$$\frac{\partial}{\partial x} f_i(a,x)|_{a=(b,x)} = \psi_i(b,x) \tag{38}$$

by virtue of (31) and (32). Using (15) and (37), we obtain

$$\int_A z\left(f\left(a,x\right)\mu\left(x,da\right)\right)=\int_B z\left(f\left(\left(b,x\right),x\right)\right)\lambda\left(x,db\right)=\int_B z\left(g\left(b,x\right)\right)\lambda\left(x,db\right)$$

for any bounded Borel function $z\left(r\right),\ r\in R^n$. This proves (7).

From (37) and (38), we find

$$\frac{\partial}{\partial x}[f_i(a,x)f_j(a,x)]|_{a=(b,x)}=g_i(b,x)\cdot\psi_j(b,x)+g_j(b,x)\psi_i(b,x). \tag{39}$$

Consequently,

$$\int_X\rho(dx)\int_A\mu(x,da)\frac{\partial}{\partial x}[f_i(a,x)f_j(a,x)]=$$

$$\int_X\rho(dx)\int_B\lambda(x,db)\frac{\partial}{\partial x}[f_i(a,x)f_j(a,x)]|_{a=(b,x)}=$$

$$\langle g_i,\psi_j\rangle+\langle g_j,\psi_i\rangle=\Phi(g_i,g_j)=\int_X\rho(dx)\frac{d}{dx}\int_A\mu(x,da)f_i(a,x)f_j(a,x),$$

where the first equality follows from (15), the second from (39), the third from (17), and the fourth from (7) (see Remark 2).

Theorem 3 is proved.

To conclude, we consider a question concerning a possibility of refining Theorem 1. Denote by $\nu\left(x,\cdot\right)$ the induced distribution $g\left(\cdot,x\right)\circ\lambda\left(x,\cdot\right)$ of demand vectors. Suppose there is a measurable space $(A,\mathcal{A})$, a probability measure μ on $\mathcal{A}$, and a mapping $f\left(a,x\right)=\left(f_1\left(a,x\right),...,f_n\left(a,x\right)\right)$ in the class $W\left(A\times X\right)$ such that

$$\int_A\max_{x\in X}\left|\frac{\partial f\left(a,x\right)}{\partial x}\right|\mu\left(da\right)<\infty, \tag{40}$$

and

$$\nu\left(x,\cdot\right)=f\left(\cdot,x\right)\circ\mu. \tag{41}$$

By virtue of (40), we have

$$\int_A \max_{x\in X}\left|\frac{\partial}{\partial x}\left[f_i(a,x)\,f_j(a,x)\right]\right|\mu(da)<\infty,\quad i,j\in\{1,...,n\}.$$

Consequently,

$$\frac{d}{dx}\int_A f_i(a,x)\,f_j(a,x)\,\mu(da)=\int_A \frac{\partial}{\partial x}\left[f_i(a,x)\,f_j(a,x)\right]\mu(da),$$

and so the metonymy hypothesis holds. A statement about the existence of A,μ, and $f(\cdot,\cdot)$ with properties (40) and (41) may be regarded as a refinement of Theorem 2 in which the measure $\mu(x,da)$ does not depend on x. Assumptions (A.1) and (A.2) imposed in this paper do not guarantee that such a statement is true. Although there always exist a space $(A,\mathcal{A})$, a probability measure μ, and an $\mathcal{A}\times\mathcal{B}(X)$-measurable mapping $f:A\times X\to R^n$ for which (41) holds[4], one cannot, in general, establish the existence of a differentiable, or even continuous, mapping $f(a,\cdot)$ with property (41). Consider a simple example.

Let $g^{(1)}\neq g^{(2)}$ be strictly positive vectors belonging to the unit simplex in R^n. Define $B=\{1,2\}$ and $g(b,x)=xg^{(b)}\ (x\in X)$. Assume that the measure $\lambda(x,\cdot)$ assigns masses $\lambda_b(x)>0$ to the points $b=1,2\in B$, where the functions $\lambda_b(x)$ are continuously differentiable and not identically constant. Then conditions (A.1) and (A.2) are satisfied. Suppose there exist $(A,\mathcal{A})$, μ, and $f(\cdot,\cdot)$ such that (41) is fulfilled and $f(a,x)$ is continuous in $x\in X$. Then the mapping $f^0(a,x)=x^{-1}f(a,x)$ is continuous

[4]One can define $(A,\mathcal{A})$ as a segment $[0,1]$ with Borel σ-algebra and μ as the Lebesgue measure. To construct $f(a,x)$, consider a mapping $\tau:R^n\to[0,1]$ which establishes a Borelean isomorphism between R^n and $[0,1]$. Define $D(r,x)=\nu\left(x,\tau^{-1}([0,r])\right)$ and put $f(a,x)=\tau^{-1}(\min\{s:D(r,x)\geq a\})$. Then $f(a,x)$ is jointly measurable in (a,x) and (41) holds .

in x, and for each $x \in X$ the measure $f^0(\cdot, x) \circ \mu$ is concentrated on $\left\{g^{(1)}, g^{(2)}\right\}$. Since $f^0(a, \cdot)$ is continuous, we obtain that the set $A^0 = \left\{a : f^0(a, x) \in \left\{g^{(1)}, g^{(2)}\right\} \text{ for all } x\right\}$ has measure 1. For each $a \in A^0$, the continuous mapping $f^0(a, x)$ takes two values and is thus a constant, $g^{(j(a))}$. Consequently, $\lambda_b(x) = \mu\left\{a \in A^0 : f^0(a, x) = g^{(b)}\right\}$ =const. A contradiction.

It would be of interest to investigate conditions under which there exist a measurable space $(A, \mathcal{A})$, a probability measure μ on $\mathcal{A}$, and a mapping $f(\cdot, \cdot) \in W(A \times X)$ possessing properties (40) and (41).

Acknowledgements

Financial support from Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged.

References

Härdle, W., W. Hildenbrand, and M. Jerison, 1991, Empirical evidence on the Law of Demand, Econometrica 59, 1525–1549.

Hildenbrand, W., 1983, On the "Law of Demand", Econometrica 51, 997–1019.

Hildenbrand, W., 1994, Market demand: theory and empirical evidence (Princeton University Press, Princeton).

Skorokhod, A.V., 1965, Studies in the theory of random processes (Addison–Wesley, Reading, Massachusetts).

Vainberg, M.M., 1973, The variational method and the method of monotone operators in the theory of nonlinear equations (Wiley, New York).

Wald, A., 1936, Über einige Gleichungssysteme der mathematischen Ökonomie, Zeitschrift für Nationalökonomie, 7, 637–670.