DISCUSSION PAPERS
Institute of Economics
University of Copenhagen

Papers by Participants in
The Birgit Grodal Symposium, Copenhagen, September 2002

03-26

On Behavioral Heterogeneity

Werner Hildenbrand and Alois Kneip

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

Birgit Grodal Symposium
Topics in Mathematical Economics

The participants in a September 2002 Workshop on *Topics in Mathematical Economics* in honor of Birgit Grodal decided to have a series of papers appear on Birgit Grodal's 60'th birthday, June 24, 2003.

The Institute of Economics suggested that the papers became Discussion Papers from the Institute.

The editor of *Economic Theory* offered to consider the papers for a special Festschrift issue of the journal with Karl Vind as Guest Editor.

This paper is one of the many papers sent to the Discussion Paper series.

Most of these papers will later also be published in a special issue of *Economic Theory*.

Tillykke Birgit

Troels Østergaard Sørensen	Karl Vind
Head of Institute	Guest Editor

On Behavioral Heterogeneity

Werner Hildenbrand and Alois Kneip

1 Introduction

It is well-known that the hypothesis of utility maximization on the individual level does not imply any identifiable properties of mean demand for a sufficiently large population H of households, in particular, the Jacobian matrix

$$\partial F(p) = \left(\partial_{p_j} F_i(p)\right)_{i,j=1,\ldots,l}$$

of mean demand

$$F(p) = \frac{1}{\#H} \sum_{h \in H} f^h\left(p, x^h\right) \tag{1}$$

has no identifiable structure, such as negative definiteness or diagonal dominance (Sonnenschein (1973), Andreu (1983) and Chiappori and Ekeland (1996)).

Consequently, in order to obtain some useful properties of mean demand, one has to make additional assumptions either on the individual utility functions or on the joint distribution of demand functions and income, that is to say, on the composition of the population of households. For the first approach see, for example, the literature on additively separable utility functions or Mitjuschin and Polterovich (1978). The distributional approach is treated in Hildenbrand (1983), Härdle, Hildenbrand and Jerison (1991), Grandmont (1992), Quah (1997), Kneip (1999) and Jerison (1999). In this paper we pursue the distributional approach.

In a very original and influential paper Grandmont (1992) showed that "increasing behavioral heterogeneity makes aggregate expenditure more independent of prices". Grandmont's definition of "behavioral heterogeneity" is based on a specific parametrization of households' demand functions (given by the α-transformation) and is defined by a *distributional assumption on the parameter.*

In this paper we want to show that Grandmont's proposition holds in a much more general setting, which does not require any parametrization of households' demand function. To prove this claim we have, of course, to define in this general setting a notion of "behavioral heterogeneity". This is done in section 2.

Let $W_i(p)$ denote the aggregate (mean) consumption expenditure ratio for commodity i of a population of households, i.e.,

$$W_i(p) = p_i \cdot F_i(p)/X \qquad\qquad i = 1, \dots, l$$

where X denotes mean income of the population.

Consider now the rate of change of $W_i(p)$ with respect to a percentage change of the price p_j, i.e.,

$$S_{ij}(p) := \partial_\lambda \left[W_i(p_1, \dots, \lambda p_j, \dots, p_l)\right]_{\lambda=1} = p_j \partial_{p_j} W_i(p).$$

With this notation we can now formulate Grandmont's main result. Let

$$W(p) = \int_{\mathbb{R}^l} w(\alpha * p, x)\, d\nu \tag{2}$$

where $w(p, x)$ is any continuously differentiable budget share function (i.e., $0 \leq w_i(p, x) \leq 1$ and $\sum_{i=1}^l w_i(p, x) = 1$), ν denotes a (probability) distribution of the parameter $\alpha \in (0, \infty)^l$ and $\alpha * p = (\alpha_1 \cdot p_1, \dots, \alpha_l \cdot p_l)$. Note that all households in the population (described by w and ν) have the same income x.

Then, $\sup_p |S_{ij}(p)|$ becomes arbitrarily small if $\beta = (\log \alpha_1, \dots, \log \alpha_l)$ is "sufficiently uniformly spread" on $\mathbb{R}^l$. For example, if β is uniformly distributed on the cube $Q = [a, b]^l$, then one obtains

$$\sup_p |S_{ij}(p)| \leq \frac{1}{(b-a)}.$$

In this paper we want to show that for every (finite) population H of households one has

$$\sup_p |S_{ij}(p)| \leq B(H) \tag{3}$$

where the bound $B(H)$ becomes arbitrarily small provided the population is sufficiently "behavioral heterogeneous", as defined in section 2. However, it is important to emphasize that the relevant question is not whether the bound

$B(H)$ is nearly zero but rather whether the bound $B(H)$ is small as compared to the mean comsumption expenditure ratio W_i, $i = 1, \ldots, l$.

In section 2 we shall associate to every population H of households an index $\gamma(H)$ between zero and one which can be interpreted as a *degree of behavioral heterogeneity* of the population. Naturally, $\gamma(H) = 0$ for a population where all households have the same demand function and the same income. The index $\gamma(H)$ is also zero for a population where all households have a Cobb-Douglas demand function even if they are different across the population. This population is heterogeneous, yet not behaviorally heterogeneous since all households have the same price-elasticities. The degree $\gamma(H)$ of behavioral heterogeneity is positive yet less than one if the households $h \in H$ react differently to price changes.

If one computes the index γ for the Grandmont model (2) one obtains an index $\gamma(w, \nu)$ arbitrarily close to one if the parameter $\beta = \log \alpha$ is sufficiently uniformly spread on $\mathbb{R}^l$. For example, if the parameter β is uniformly distributed on the cube $Q = [a, b]^l$ one obtains $\gamma \geq 1 - \frac{1}{(b-a)} \cdot c$, where the coefficient c is determined by the generating budget share function w.

The degree $\gamma(H)$ of behavioral heterogeneity is used to determine the bound $B(H)$ in the inequality (3). For example, consider the special case of a population $\left\{w^h, x^h\right\}_{h \in H}$ with $x^h = x$ and $\sup_p p_j |\partial_{p_j} w_i^h\left(p, x^h\right)| = d_{ij}$ for all $h \in H$. Note that Grandmont's model satisfies this assumption. Then we obtain (Proposition 3)

$$\sup_p |S_{ij}(p)| \leq (1 - \gamma(H))\, d_{ij}.$$

For a general population the relation between the degree of behavioral heterogeneity $\gamma(H)$ and the bound $B(H)$ is more complicated and is given in Proposition 4.

Obviously, a bound on $|S_{ij}(p)|$ implies a structural property of the Jacobian $\partial F(p)$ of mean demand, since

$$D_p \partial F(p)\, D_p = X\left(-D_{W(p)} + S(p)\right) \tag{4}$$

where D_p and $D_{W(p)}$ are diagonal matrices with $p_1, \ldots, p_l$ and $W_1(p), \ldots, W_l(p)$ on the diagonal. Consequently, if $|S_{ij}(p)|$ *is small as compared with* $W_i(p)$, then the diagonal of the Jacobian matrix $\partial F(p)$ becomes negative and dominant. In particular, if for every i, $W_i(p)$ is greater than $\sum_{j=1}^{l} |S_{ij}(p)|$ and $\sum_{j=1}^{l} |S_{ji}(p)|$

then the Jacobian matrix $\partial F(p)$ is negative definite and mean demand $F(p)$ satisfies the Law of Demand, i.e.,

$$(p-q)\cdot(F(p)-F(q))<0.$$

Finally we remark that a bound on $|S_{ij}(p)|$ implies a restriction on the magnitude of price-elasticities of mean demand. Let $E_{ij}(p)$ denote the price-elasticity of mean demand for commodity i with respect to the price p_j. Then one obtains

$$\begin{aligned} E_{ii}(p) &= -1+S_{ii}(p)/W_i(p) \\ E_{ij}(p) &= S_{ij}(p)/W_i(p) \qquad \text{if } i\neq j. \end{aligned}$$

Consequently, a low bound for $|S_{ij}(p)|$ implies strong restrictions on the matrix $E(p)$ of price-elasticities.

In section 3 we shall use the estimates of price-elasticities obtained in Blundell et al. (1993) in order to check whether the implication of the hypothesis of "behavioral heterogeneity" is compatible with empirical facts.

2 Behavioral Heterogeneity

Notation: A population H of households $h\in H$ is defined by $\{f^h,x^h\}_{h\in H}$, where $(p,x)\longmapsto f^h(p,x)\in\mathbb{R}^l_+$ denotes the demand function and $x^h>0$ the income of household h. The price vector $p\in\mathbb{P}^l=(0,\infty)^l$.

Mean demand $F(p)$ of the population H is defined by

$$F(p):=\frac{1}{\#H}\sum_{h\in H}f^h\left(p,x^h\right).$$

The *consumption expenditure ratio of household* h for commodity i (also called the budget share) and the *aggregate consumption expenditure ratio* are defined by

$$w_i^h(p,x):=\frac{1}{x}p_i\cdot f_i^h(p,x) \qquad i=1,\ldots,l$$

and

$$W_i(p) := \frac{1}{X} p_i F_i(p),$$

where X denotes mean income $X = \frac{1}{\#H} \sum_{h \in H} x^h$ of the population.

Thus,

$$W_i(p) = \frac{1}{\#H} \sum_{h \in H} \frac{x^h}{X} w_i^h\left(p, x^h\right). \tag{1}$$

Let s_{ij}^h and S_{ij} denote the rate of change of w_i^h and W_i, respectively, with respect to a percentage change of the price p_j, i.e.,

$$s_{ij}^h(p,x) := \partial_\lambda w_i^h(p_1, \ldots, \lambda p_j, \ldots, p_l)\,|_{\lambda=1} = p_j \partial_{p_j} w_i^h(p,x)$$

$$S_{ij}(p) := \partial_\lambda W_i(p_1, \ldots, \lambda p_j, \ldots, p_l)\,|_{\lambda=1} = p_j \partial_{p_j} W_i(p).$$

We want to find an upper bound for $\sup_p |S_{ij}(p)|$. If one just knows for every household the income x^h and $d_{ij}^h := \sup_p |s_{ij}^h(p, x^h)|$ then it follows from (1) that the best possible upper bound is given by

$$\sup_p |S_{ij}(p)| \leq \frac{1}{\#H} \sum_{h \in H} \frac{x^h}{X} d_{ij}^h =: D_{ij}. \tag{2}$$

We remark that d_{ij}^h does not depend on the income level x^h if the budget share function w^h is homogeneous of degree zero in (p,x).

It is our goal to improve upon inequality (2). To achieve this, one needs some information about the characteristics of the population. The question now is: which feature of the population H leads to an improvement of inequality (2)?

We shall associate to every population H of households an index $\gamma(H)$ with $0 \leq \gamma(H) \leq 1 - \frac{1}{\#H}$, which will be interpreted as a *degree of behavioral heterogeneity* of the population H. The index $\gamma(H)$ is used to improve upon inequality (1).

The following assumptions on the budget share function $w(p,x)$ are made:

Assumption 1: $0 \leq w_i(p,x) \leq 1$

This follows from the budget restriction $p \cdot f(p,x) \leq x$.

Assumption 2: $w(p,x)$ is continuously differentiable in $p \gg 0$ and $x > 0$ and

$$\sup_p |p_j \partial_{p_j} w_i(p,x)| < \infty.$$

Note that $\sup_p |\partial_{p_j} w_i(p,x)|$ is typically not finite. For this reason we consider the rate of change of $w_i(p,x)$ for a percentage change of the j-th price.

Assumption 3: there is an integer m such that for all budget share functions that we shall consider and for every $i,j \in \{1,\ldots,l\}$ and every $\overline{p} \in \mathbb{P}^l$ the derivative of the function

$$p_j \longmapsto w_i(\overline{p}_1,\ldots,p_j,\ldots,\overline{p}_l,x)$$

changes its sign at most m times.[1]

Let

$$A_{ij}^{\varepsilon}(w,x) := \left\{p \in \mathbb{P}^l \middle| \;\; p_j |\partial_{p_j} w_i(p,x)| \geq \varepsilon \cdot d_{ij}(w,x)\right\} \tag{3}$$

where $d_{ij}(w,x) = \sup_p \left\{p_j |\partial_{p_j} w_i(p,x)|\right\}$. That is to say, the set $A_{ij}^{\varepsilon}(w,x)$ is the domain of price vectors p for which the absolute value of the rate of change of the budget share w_i for commodity i with respect to a percentage change of the price p_j is larger than or equal to $\varepsilon \cdot d_{ij}(w,x)$.

Note that $A_{ij}^{0}(w,x) = \mathbb{P}^l$, $A_{ij}^{\varepsilon}(w,x) \neq \emptyset$ and is decreasing in ε for $0 < \varepsilon < 1$, and $A_{ij}^{\varepsilon}(w,x) = \emptyset$ for $\varepsilon > 1$.

The sets $A_{ij}^{\varepsilon}(w,x)$ will play a crucial role in defining the notion of "behavioral heterogeneity". Therefore it is important to emphasize that the set $A_{ij}^{\varepsilon}(w,x)$ depends in an essential way on the budget share function w and the income level x. To show this dependence it turns out that it is convenient to consider log-prices

$$q = (q_1,\ldots,q_l) = (\log p_1,\ldots,\log p_l) = \log p.$$

Then one defines

$$B_{ij}^{\varepsilon}(w,x) := \left\{q \in \mathbb{R}^l \middle| \;\; |\partial_{q_j} w_i(\exp(q),x)| \geq \varepsilon \cdot d_{ij}(w,x)\right\}. \tag{4}$$

[1] Without loss of economic content one might even assume that $m = 0$, that is to say, the function $p_j \longmapsto s_{ij}^h(\overline{p}_1,\ldots,p_j,\ldots,\overline{p}_l,x^h)$ is either non-negative or non-positive.

One easily verifies that $p \in A_{ij}^{\varepsilon}(w, x)$ if and only if $\log p \in B_{ij}^{\varepsilon}(w, x)$.

Proposition 1 *If the budget share function w (or, equivalently, the demand function f) is homogeneous of degree zero in (p, x), then*

$$\begin{aligned} B_{ij}^{\varepsilon}(w, x) &= B_{ij}^{\varepsilon}(w, 1) + \log x \cdot \mathbf{1} \\ &= \left\{ (q_1 + \log x, \ldots, q_l + \log x) \,|\, q \in B_{ij}^{\varepsilon}(w, 1) \right\}. \end{aligned}$$

If, in addition, $w(p, x)$ does not depend on the income level x (or equivalently, the demand function $f(p, x)$ is linear in x), then the set $B_{ij}^{\varepsilon}(w, x)$ does not depend on x and $q \in B_{ij}^{\varepsilon}(w)$ implies $q + \lambda \mathbf{1} \in B_{ij}^{\varepsilon}(w)$ for all $\lambda \in \mathbb{R}$.

Proof: First we remark that homogeneity of w in (p, x) implies that $\sup_p \left| p_j \partial_{p_j} w_i(p, x) \right| = d_{ij}(w, x)$ does not depend on the income level x. Homogeneity of w in (p, x) implies $w_i(p, x) = w_i\left(\frac{1}{x} p, 1\right)$, and hence

$$\partial_\lambda w_i(p_1, \ldots, \lambda p_j, \ldots, p_l, x)\,|_{\lambda=1} = \partial_\lambda w_i(p_1/x, \ldots, \lambda p_j/x, \ldots, p_L/x, 1)\,|_{\lambda=1}.$$

Since $d_{ij}(w, x)$ does not depend on x we obtain that $p \in A_{ij}^{\varepsilon}(w, x)$ if and only if $\frac{1}{x} p \in A_{ij}^{\varepsilon}(w, 1)$, and hence $q \in B_{ij}^{\varepsilon}(w, x)$ if and only if $\log q - \log x \cdot 1 \in B_{ij}^{\varepsilon}(w, 1)$, which proves the first part of Proposition 1.

If $w(p, x)$ does not depend on x then, by definition of the set $B_{ij}^{\varepsilon}(w, x)$, this set does not depend on x and we write $B_{ij}^{\varepsilon}(w)$. Hence, if $q \in B_{ij}^{\varepsilon}(w)$, then the first part of Proposition 1 implies $q - \log x \cdot 1 \in B_{ij}^{\varepsilon}(w)$ for every $x > 0$. Q.E.D.

The dependence of the sets $A_{ij}^{\varepsilon}(w, x)$ and $B_{ij}^{\varepsilon}(w, x)$ on the budget share function w is complex. We just give two examples.

Example 1: Consider two budget share functions w and w^{α} which are linked by

$$w^{\alpha}(p, x) = w(\alpha * p, x) \tag{α}$$

for some $\alpha \in \mathbb{P}^l$ where $\alpha * p = (\alpha_1 \cdot p_1, \ldots, \alpha_l \cdot p_l)$. The corresponding demand function f and f^{α} then satisfy

$$f^{\alpha}(p, x) = \alpha * f(\alpha * p, x).$$

This transformation of demand functions has been considered by Mas-Colell and Neuefeind (1977), E. Dierker, H. Dierker and Trockel (1984) and Grandmont (1992).

One easily shows that (α) implies

$$B^{\varepsilon}_{ij}\left(w^{\alpha},x\right)=B^{\varepsilon}_{ij}\left(w,x\right)-\log\alpha \quad ,$$

that is to say, $B^{\varepsilon}_{ij}\left(\omega^{\alpha},x\right)$ is just a translation of $B^{\varepsilon}_{ij}\left(\omega,x\right)$ by the vector $\left(\log\alpha_1,\ldots,\log\alpha_l\right)$. Indeed, we first observe that $d_{ij}\left(w,x\right)=d_{ij}\left(w^{\alpha},x\right)$. Let $q\in B^{\varepsilon}_{ij}\left(w^{\alpha},x\right)$, i.e., $\left|\partial_{q_j}w^{\alpha}_i\left(\exp q,x\right)\right|\geq\varepsilon d_{ij}\left(w^{\alpha},x\right)$. Since $\partial_{q_j}w^{\alpha}_i\left(\exp q,x\right)=\partial_{q_j}w_i\left(\alpha*\exp q,x\right)=\partial_{q_j}w_i\left(\exp\left(\log\alpha+q\right),x\right)$ it follows that $\log\alpha+q\in B^{\varepsilon}\left(w,x\right)$. Thus we showed that $B^{\varepsilon}_{ij}\left(w^{\alpha},x\right)\subset B^{\varepsilon}\left(w,x\right)-\log\alpha$. The opposite inclusion is shown analogously.

Example 2: CES demand functions

We only consider the case of two commodities. The CES budget share functions are independent of income and are defined by

$$\begin{aligned}\omega_1\left(p_1,p_2\right) &= \frac{a^{\sigma}p_1^{1-\sigma}}{a^{\sigma}p_1^{1-\sigma}+\left(1-a\right)^{\sigma}p_2^{1-\sigma}}\\ \omega_2\left(p_1,p_2\right) &= \frac{\left(1-a\right)^{\sigma}p_2^{1-\sigma}}{a^{\sigma}p_1^{1-\sigma}+\left(1-a\right)^{\sigma}p_2^{1-\sigma}}\end{aligned}$$

where (a,σ) are parameters with $0\leq a\leq 1$ and $\sigma>0$.

The functions $s_{ij}\left(p_1,p_2\right)=p_j\partial_{p_j}\omega_i\left(p_1,p_2\right)$ are either everywhere positive or negative, depending on the parameter values (a,σ). Hence, assumption 3 on budget share functions is satisfied for $m=0$.

Figure 1 shows the graph of $s_{11}\left(\cdot\right)$ as a function of $\left(\log p_1,\log p_2\right)$ for the parameter values $(a,\sigma)=(0.95,2)$, $(0.5,2)$, $(0.05,2)$ and $(0.5,0.1)$.

It follows from Proposition 1 that the set $B^{\varepsilon}_{11}\left(a,\sigma\right)$ is a strip parallel to the diagonal in $\mathbb{R}^2$. Thus,

$$B^{\varepsilon}_{11}\left(a,\sigma\right)=\left\{\left(u,v\right)\in\mathbb{R}^2\middle|\ \ v-u\in\left[z-b_l,z+b_r\right]\right\}.$$

One can compute the interval $\left[z-b_l,z+b_r\right]$ and show that the length of this interval only depends on the parameter σ and ε, while the location additionally depends on the parameter a.

In Figure 2 are plotted the sets $B^{\varepsilon}_{11}\left(a,\sigma\right)\cap Q$ for $\varepsilon=0.7$ and the parameter values $(a,\sigma)=(0.95,2.5)$, $(0.5,2)$ and $(0.05,1.5)$, and the cube $Q=\left[-20,20\right]^2$. By a closer look at Figure 2 we recognize that for all three functions the corresponding set of prices $\log p\in B^{\varepsilon}_{11}\left(a,\sigma\right)$ only covers a small fraction of the whole cube $\left[-20,20\right]^2$. This is not a specific property of CES-functions, but holds for all budget share functions w satisfying our assumptions. Indeed, we shall show that for $\varepsilon > 0$ the set $B_{ij}^{\varepsilon}(w, x)$ is sparse in $\mathbb{R}^l$, that is to say, for any cube $Q = [a, b]^l$ in $\mathbb{R}^l$

$$\frac{\lambda^l\left(B_{ij}^{\varepsilon}(w, x) \cap Q\right)}{\lambda^l(Q)} \underset{(b-a)\to\infty}{\longrightarrow} 0$$

where λ^l denotes the Lebesgue measure on $\mathbb{R}^l$.

Proposition 2 *If $d_{ij}(w, x) > 0$ and $\varepsilon > 0$ then for every $\overline{q} \in \mathbb{R}^l$*

$$\lambda^1\left\{q_j \in \mathbb{R} \middle| \quad (\overline{q}_1, \ldots, q_j, \ldots, \overline{q}_l) \in B_{ij}^{\varepsilon}(w, x)\right\} \leq \frac{m+1}{\varepsilon \cdot d_{ij}(w, x)}$$

Corollary: *The set $B_{ij}^{\varepsilon}(w, x)$ is sparse for $\varepsilon > 0$*, more specifically,

$$\frac{\lambda^l\left(B_{ij}^{\varepsilon}(w, x) \cap Q\right)}{\lambda^l(Q)} \leq \frac{m+1}{\varepsilon \cdot d_{ij}(w, x)(b-a)}.$$

Proof of Proposition 2: Let $v(\xi) := w_i\left(\exp \overline{q}_1, \ldots, \exp \overline{q}_{j-1}, \exp \xi, \exp \overline{q}_{j+1}, \ldots, x\right)$. By definition of the set $B_{ij}^{\varepsilon}(w, x)$ we obtain

$$\left\{q_j \in \mathbb{R} \middle| \quad (\overline{q}_1, \ldots, q_j, \ldots, \overline{q}_l) \in B_{ij}^{\varepsilon}(w, x)\right\} = \{\xi \in \mathbb{R} | \quad |v'(\xi)| \geq \varepsilon d_{ij}(w, x)\} =: C.$$

Assumption 3 on budget share functions implies that there are $m+1$ intervals, $I_1 = (-\infty, z_1), \ldots, I_n = (z_{n-1}, z_n), \ldots, I_{m+1} = (z_m, \infty)$ such that the function v is monotone on every interval. Hence $\int_{I_n} |v'(\xi)| \, d\xi = \left|\int_{I_n} v'(\xi) \, d\xi\right|$. Since $0 \leq v(\xi) \leq 1$ one obtains

$$\left|\int_{I_n} v'(\xi) \, d\xi\right| = |v(z_n) - v(z_{n-1})| \leq 1.$$

Note that $v(z_0) = \lim_{\xi \longrightarrow -\infty} v(\xi)$ and $v(z_{m+1}) = \lim_{\xi \longrightarrow \infty} v(\xi)$ exist. Consequently,

$$1 \geq \int_{I_n} |v'(\xi)| \, d\xi \geq \int_{C \cap I_n} |v'(\xi)| \, d\xi \geq \varepsilon d_{ij}(w, x) \lambda^1(C \cap I_n),$$

which implies Proposition 2. Q.E.D.

Our discussion of properties of the sets $B_{ij}^{\varepsilon}\left(w^h, x^h\right)$ allows to draw the following conclusion: The sets $B_{ij}^{\varepsilon}\left(w^h, x^h\right)$ are sparse, and their exact location in $\mathbb{R}^l$

depends crucially on the households' characteristics $\left(w^h, x^h\right)$. If (w^1, x^1) is close to (w^2, x^2), then $B^{\varepsilon}_{ij}(w^1, x^1)$ and $B^{\varepsilon}_{ij}(w^2, x^2)$ are to be found in similar regions. On the other hand, if there are substantial structural differences between two budget share functions, then the corresponding sets B^{ε}_{ij} do not intersect. This has already been illustrated by the CES example (see Figure 2). In the context of example 1 we can infer from the sparseness of $B^{\varepsilon}_{ij}(w, x)$ that $B^{\varepsilon}_{ij}(w, x) \cap B^{\varepsilon}_{ij}(w^{\alpha}, x) = \emptyset$ if the shift introduced by α is sufficiently large.

The intuitive notion of "behavioral heterogeneity" of a population H of households - however it is made precise - undoubtedly implies that households' characteristics $\left(w^h, x^h\right)$ must be different across the population. This alone, however, is not sufficient, and the above arguments open a way to quantify the structurally important differences. "Behavioral heterogeneity" should imply that the sets $B^{\varepsilon}_{ij}\left(w^h, x^h\right)$ and, hence, $A^{\varepsilon}_{ij}\left(w^h, x^h\right)$ are different in the sense that they are "located in different regions" in $\mathbb{R}^l$, which could be made precise by requiring that these sets possess only few intersections. In terms of the sets $A^{\varepsilon}_{ij}(w, x)$ we now define the *intersection frequency:* for every $\varepsilon \geq 0$ we define

$$I^{\varepsilon}_{ij}(p) := \frac{1}{\#H} \# \left\{ h \in H \mid \; p \in A^{\varepsilon}_{ij}\left(w^h, x^h\right) \right\} = \frac{1}{\#H} \# \left\{ h \in H \mid \; \log p \in B^{\varepsilon}_{ij}\left(w^h, x^h\right) \right\}.$$

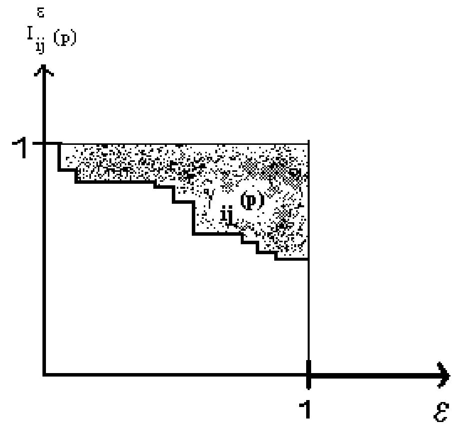

Figure 1:

Obviously, $I_{ij}^{0}(p) = 1$ and $I_{ij}^{\varepsilon}(p)$ is a decreasing step-function in ε.

A behaviorally heterogeneous population will now be characterized by the property that for every price vector p the intersection frequency $I_{ij}^{\varepsilon}(p)$ is small. Since we do not want to emphasize a particular value of ε we shall consider the shaded area $\gamma_{ij}(p)$ in Figure 3, i.e.,

$$\gamma_{ij}(p) := 1 - \int_{0}^{1} I_{ij}^{\varepsilon}(p)\, d\varepsilon$$

and define the *degree of behavioral heterogeneity* with respect to i, j by

$$\gamma_{ij} := \inf_{p} \gamma_{ij}(p).$$

Definition: The *degree of behavioral heterogeneity of the population* H is defined by

$$\inf_{p,i,j} \gamma_{ij}(p) =: \gamma(H).$$

The minimal degree of behavioral heterogeneity, i.e., $\gamma(H) = 0$, is attained for a population H if

- all households have the same demand function and the same income
- all households have the same demand function which is linear in income
- all households have a Cobb-Douglas demand function which can be different across the population.

Obviously, $\gamma(H) \leq 1 - \frac{1}{\#H} < 1$ for every population H.

If there is an ε such that the sets $B_{ij}^{\varepsilon}\left(w^h, x^h\right)$ are *disjoint* (see example 2) then $I_{ij}^{\varepsilon}(p) \leq 1/\#H$ for every p and hence $\gamma_{ij} \geq (1-\varepsilon)\left(1 - \frac{1}{\#H}\right)$.

Proposition 3 *For every population H with $d_{ij}^h = d_{ij}$ and $x^h = x$ one obtains*

$$\sup_p |S_{ij}(p)| \leq (1 - \gamma(H))\, d_{ij}.$$

Proof:
Define $G_{ij}^{\varepsilon}(p) := 1 - I_{ij}^{\varepsilon}(p)$. Then

$$G_{ij}^{\varepsilon}(p) = \frac{1}{\#H}\#\left\{h \in H \mid \ |s_{ij}^h(p,x)| < \varepsilon d_{ij}\right\}.$$

Thus, $G_{ij}^{\varepsilon}(p)$ is the cumulative distribution function of $|s_{ij}^h(p,x)|/d_{ij}$. Consequently,

$$\begin{aligned} |S_{ij}(p)| &\leq \frac{1}{\#H}\sum_{h\in H}\left|s_{ij}^h(p,x)\right| = d_{ij}\cdot\frac{1}{\#H}\sum_{h\in H}\left|s_{ij}^h(p,x)\right|/d_{ij} \\ &= d_{ij}\cdot\int \varepsilon\, dG_{ij}^{\varepsilon}(p). \end{aligned}$$

Since $\int \varepsilon\, dG_{ij}^{\varepsilon}(p) = 1 - \gamma_{ij}(p)$ we obtain

$$|S_{ij}(p)| \leq \left(1 - \gamma_{ij}(p)\right) d_{ij}.$$

We remark that one obtains

$$\sup_p |S_{ij}(p)| = (1 - \gamma(H))\, d_{ij}$$

if for all households $h \in H$ and all price vectors p either $s_{ij}^h\left(p, x^h\right) \geq 0$ or $s_{ij}^h\left(p, x^h\right) \leq 0$.

Example 3: (Grandmont (1992))

Recall from example 1 that $B_{ij}^{\varepsilon}(w^{\alpha}, x) = B_{ij}^{\varepsilon}(w^1, x) - \log\alpha,\ \alpha \in \mathbb{R}_+^l$. The population is described by the budget-share function w^1, the common income x and the distribution ν of $\log\alpha$ on $\mathbb{R}^l$. The intersection frequency is defined by

$$\begin{aligned} I_{ij}^{\varepsilon}(p) &= \nu\left\{\log\alpha \in \mathbb{R}^l \mid \ \log p \in B_{ij}^{\varepsilon}(w^{\alpha}, x)\right\} \\ &= \nu\left\{\log\alpha \in \mathbb{R}^l \mid \ \log\alpha \in B_{ij}^{\varepsilon}\left(w^1, x\right) - \log p\right\} = \nu\left\{B_{ij}^{\varepsilon}\left(w^1, x\right) - \log p\right\}. \end{aligned}$$

Since the set $B^{\varepsilon}_{ij}(w^1, x)$, and hence its translate by $\log p$, is sparse it follows that the intersection frequency $I^{\varepsilon}_{ij}(p)$ is arbitrarily small if the distribution ν is sufficiently uniformly spread. For example, if ν is the uniform distribution on the cube $Q = [a, b]^L$ then we obtain from Proposition 2 that

$$I^{\varepsilon}_{ij}(p) \leq \frac{m+1}{\varepsilon d_{ij}(b-a)}.$$

Finally we consider a general population H. The goal is to improve upon inequality (2).

It follows from (1) that

$$\begin{aligned} |S_{ij}(p)| &\leq \frac{1}{\#H} \sum_{h \in H} \left| s^h_{ij}\left(p, x^h\right) \right| x^h / X \\ &= \frac{1}{\#H} \sum_{h \in H} \frac{\left| s^h_{ij}\left(p, x^h\right) \right|}{d^h_{ij}} \cdot \frac{d^h_{ij} x^h}{X}. \end{aligned}$$

Since $\frac{1}{\#H} \sum_{h \in H} \frac{\left| s^h_{ij}\left(p, x^h\right) \right|}{d^h_{ij}} = 1 - \gamma_{ij}(p)$ and $\frac{1}{\#H} \sum_{h \in H} \frac{d^h_{ij} x^h}{X} =: D_{ij}$, we obtain

$$|S_{ij}(p)| \leq \left(1 - \gamma_{ij}(p)\right) D_{ij} + cov_H \left(\frac{\left| s^h_{ij}\left(p, x^h\right) \right|}{d^h_{ij}}, \frac{d^h_{ij} x^h}{X} \right). \tag{5}$$

Hence, if the covariance were negative, we would obtain a result that is analogous to Proposition 3, namely

$$\sup_p |S_{ij}(p)| \leq (1 - \gamma(H)) D_{ij}.$$

Of course, the above covariance need not to be negative. Yet one can show that it becomes small for sufficiently large $\gamma(H)$.

Proposition 4 *For every population H of households with $\gamma(H) \geq \frac{1}{2}$ follows*

$$\sup_p |S_{ij}(p)| \leq (1 - \gamma(H)) D_{ij} + [\gamma(H)(1 - \gamma(H))]^{\frac{1}{2}} \cdot \frac{1}{X} \cdot \left[var_H d^h_{ij} x^h \right]^{\frac{1}{2}}.$$

Proof: By Cauchy-Schwarz inequality we have

$$\left(cov_H \left(\frac{\left| s^h_{ij}\left(p, x^h\right) \right|}{d^h_{ij}}, \frac{d^h_{ij} x^h}{X} \right) \right)^2 \leq var_H \left(\frac{\left| s^h_{ij}\left(p, x^h\right) \right|}{d^h_{ij}} \right) \cdot var_H \left(\frac{d^h_{ij} \cdot x^h}{X} \right). \tag{6}$$

Let $z_h := \frac{\left|s_{ij}^h(p,x^h)\right|}{d_{ij}^h}$. Then $0 \le z_h \le 1$ for $h \in H$ and $mean_H(z_h) = 1 - \gamma_{ij}(p)$. Since $var_H(z_h) = mean_H(z_h^2) - (mean_H(z_h))^2$, we obtain

$$var_H(z_h) \le \left(1 - \gamma_{ij}(p)\right) - \left(1 - \gamma_{ij}(p)\right)^2 = \gamma_{ij}(p)\left(1 - \gamma_{ij}(p)\right).$$

The function $\xi(1-\xi)$ is decreasing on the interval $\left[\frac{1}{2}, 1\right]$. Consequently we obtain for $\gamma(H) = \inf_{i,j,p} \gamma_{ij}(p) \ge \frac{1}{2}$ that $var_H(z_h) \le \gamma(H)(1-\gamma(H))$. Hence, from (5) and (6) we obtain the claimed inequality of Proposition 4.

Example 2 (continued): In Figure 2 we have considered the set $B_{11}^{\varepsilon}(a,\sigma)$ for a population of three households with CES budget share functions and the parameters

$$(a,\sigma) = (0.95, 2.5), (0.5, 2), (0.05, 1.5).$$

Let us additionally assume that all three households have the same income $x = x^h$ (note that $d_{ij}^1 \ne d_{ij}^2 \ne d_{ij}^3$).

A numerical calculation yields:

- $\gamma(H) = 0.650$; this is close to the upperbound for the degree of behavioral heterogeneity $1 - \frac{1}{\#H} = \frac{2}{3}$.
- $\sup_p S_{ij}(p) = 0.127$ for all i, j, which is clearly smaller than $D_{ij} = 0.250$ (see the inequality (2) in section 2).
- the bound in Proposition 4 is equal to 0.136. Note that the covariance term in (5) is positive since $(1 - \gamma(H)) D_{ij} = 0.087$.

Remark 1: The mean $|S_{ij}(p)| = \left|\frac{1}{\#H}\sum_{h\in H}\frac{x^h}{X}s_{ij}^h(p,x^h)\right|$ can be expected to be smaller than $\frac{1}{\#H}\sum_{h\in H}\frac{x^h}{X}\left|s_{ij}^h(p,x^h)\right|$, and hence, smaller than the bound given in Proposition 4 if there is a *balancing sign-effect across the population*, that is to say, some $s_{ij}^h(p,x^h)$ are positive while others are negative.

The budget identity and homogeneity of the budget share function $w(p,x)$ imply that

$$\sum_{i=1}^{l}\sum_{j=1}^{l} s_{ij}^h(p,x^h) = 0.$$

Thus, if $s_{ij}^h(p,x^h)$ is not zero for all i,j then, for a given household h, some $s_{ij}^h(p,x^h)$ are positive and others are negative. One might argue that "behavioral heterogeneity" of a population should imply that, for given i,j, the sign of

$s_{ij}^h\left(p,x^h\right)$ alternates across the population. The sign of $s_{ij}^h\left(p,x^h\right)$ changes across the population if, for $i=j$, the own price-elasticity of demand for commodity i is spread around minus one and if, for $i\neq j$, the cross price elasticity is spread around zero. Why should this prevail for every commodity? We believe that to a certain extent a balancing sign-effect always prevails, yet this balancing effect is not sufficiently general and strong in order to base on it alone the desired conclusion of a small $|S_{ij}(p)|$.

Remark 2: One might argue that the structure of the Jacobian matrix of demand depends on the level of commodity aggregation. In the simple case of Hicks-Leontief (composite-commodity) aggregation one can show that *changing signs* of $S_{ij}\left(p\right)$ *across the elementary commodities* i,j has a balancing effect which goes in favour of diagonal dominance of the Jacobian matrix of the demand system for composite-commodities.

3 Empirical Results

We have already mentioned in the introduction that there is a relation between the derivatives $S_{ij}\left(p\right)$ and the price-elasticities of mean demand. In matrix notation we obtain

$$-D_{W(p)}+S\left(p\right)=D_{W(p)}E\left(p\right).$$

Recall that $E\left(p\right)$ is the matrix of price-elasticities, while $D_{W(p)}$ is the diagonal matrix with $W_1\left(p\right),\ldots,W_l\left(p\right)$ on the diagonal. We have shown in Section 2 that in the case of extreme heterogeneity all elements of $S\left(p\right)$ will be very small, and hence $-D_{W(p)}\approx D_{W(p)}E\left(p\right)$. Consequently, in this extreme situation own price elasticities will be close to -1, while cross price elasticities will be close to 0.

On the other hand, we already argued that it might not be reasonable to expect an extremely high degree of heterogeneity. The actual degree depends on the structure of the underlying population. It is thus an interesting **empirical** question to check whether for **existing** populations the matrix $-D_{W(p)}+S\left(p\right)$ possesses the structure to be assumed in the presence of a reasonably high degree of heterogeneity: off-diagonal elements are small; all diagonal elements are negative, and their absolute values are considerably larger than those of the off-diagonal ones.

In the econometric literature one can find quite a number of empirical studies providing estimates of price elasticities which can in principle be used to answer this question. However, estimating price elasticities as required by our approach is not easy. Most studies consider aggregate time series W_{it}, X_t and p_{it} over different time periods (usually years) $t = 1, 2, \ldots$. Time changes of the W_{it} are modelled as functions of X_t and $p_{1t}, \ldots, p_{nt}$, and estimates of price elasticities are then obtained from the resulting model coefficients. In order to have sufficient data to ensure reliable estimates, it is necessary to rely on sufficiently long time series based on many periods. In view of the above definition of price elasticities this is not without problems, since it is implicitely required that only prices change, while the composition of the population remains invariant. This is certainly not fulfilled for data stemming from a wide range of different years, and consequently the resulting estimates of price elasticities may not be interpretable as required in this paper.

To our knowledge, there exists only one published work in the literature which explicitely deals with these problems. This is the well-known and very extensive empirical study presented by Blundell et al. (1993). The authors use data from the British Family Expenditure Survey (FES) over 15 years (1970-1984). In order to obtain sufficiently long time series they use monthly data.

Blundell et al. (1993) provide estimates of price elasticities based on two different models. The first one is referred to as "micro model", and it is based on a parametric model of individual budget share functions. The approach uses the fact that the FES contains cross-section data providing information about income and other important household characteristics. Model parameters are allowed to vary in dependence of individual household characteristics which results in different budget share functions for different household types. Estimates of prices elasticities are then obtained by evaluating the model at average shares and household characteristics.

Conceptually, the resulting elasticities of an "average individual" do not exactly correspond to our definition of the aggregate elasticities determining $E(p)$. We thus concentrate on the estimates of price-elasticities obtained by the second approach presented in Blundell et al. (1993), which is called the "macro model". The authors present a sophisticated model of mean budget shares as a function of prices and mean income. The model incorporates some aggregation factors which are able to partially compensate distributional changes.

Using this "macro model", Table 4c Blundell et al. (1993) provides the estimates of price elasticities for 6 commodity aggregates: food, alcohol, fuel, clothing, transport and services. The corresponding matrix $D_{W(p)}$ of mean budget share can be approximated from the average shares given in their table 2b. We have used these values to calculate a corresponding estimate of the matrix $-D_{W(p)} + S(p) = D_{W(p)} E(p)$ is presented in our Table 1.

Remark: Throughout the paper of Blundell et al. (1993) mean budget shares $W_{it}^* = \frac{1}{N} \sum_h w_h^h (p, x^h)$ are used instead of $W_{it} = C_{it}(p) / X_t$. Since usually $W_{it}^* \neq W_{it}$ this constitutes some deviation from our definitions. It might also be noted that the qualitative conclusions to be drawn from Table 1 rest valid if elasticities from the "micro model" are applied. In fact, Blundell et al. (1993) show that most of the differences in the estimated elasticities from these models are statistically not significant.

One can infer from Table 1 that all diagonal elements of $D_{W(p)} E(p)$ are negative. Furthermore, the off-diagonal elements are indeed considerably smaller in absolute value than the diagonal ones. One might say that a matrix with a similar qualitative structure is to be expected for a population with a high, but not extremely high, degree of heterogeneity.

One even recognizes that, with the exception of the commodity fuel, the diagonal elements seem to dominate the off-diagonal ones in the same row or column. One might thus ask whether even the hypothesis of diagonal dominance is acceptable. For simplicity, in order to avoid a separate discussion of sums of rows and columns we have considered this question by relying on the symmetrized matrix

$$Q(p) = D_{W(p)} E(p) + E(p)^T D_{W(p)}.$$

Diagonal dominance can now be checked by computing the differences $Q(p)_{ii} - \sum_{j \neq i} |Q_{ij}(p)|$ for each row $i = 1, \ldots, 6$. The resulting differences are reported in Table 2. Since the estimation error is by no means negligible, the table also provides rough approximations of the standard error of the estimated differences and of the corresponding lower and upper confidence bound. The indicated errors must be considered as very crude approximations, since they are calculated only from the standard errors of the elasticity estimates given in Table 4C of Blundell et al. (1993). Correlations between elasticities estimates are not provided and thus have to be ignored.

We see that five out of the six differences are negative. Positivity of the difference for the commodity aggregate "fuel" does not seem to be significant (recall, however, that we use a very crude approximation of the standard error).

Table 1: Estimates of $-W(p)+S(p)$

	Food	Alcohol	Fuel	Clothing	Transport	Services
Food	-0.211	0.009	-0.005	-0.005	0.017	0.034
Alcohol	0.007	-0.087	0.029	0.019	0.039	0.014
Fuel	-0.005	0.029	-0.046	0.000	-0.034	-0.007
Clothing	0.066	-0.042	-0.044	-0.122	0.019	-0.003
Transport	0.020	-0.007	-0.015	-0.027	-0.214	-0.018
Services	0.010	0.004	-0.004	0.010	-0.023	-0.131

Table 2: Differences

	diff	std	low.bound	upp.bound
Food	-0.252	0.074	-0.397	-0.108
Alcohol	-0.036	0.051	-0.135	0.064
Fuel	0.079	0.049	-0.017	0.176
Clothing	-0.101	0.059	-0.217	0.016
Transport	-0.259	0.095	-0.446	-0.072
Services	-0.149	0.056	-0.258	-0.040

References

Blundell, R., Pashardes, P., and Weber, G. (1993). What do we learn about consumer demand patterns from micro data? *The American Economic Review*, 83, 570–597.

Chiappori, P. A. and Ekeland, I. (1996). Problemes dt'agregation en theorie du consommateur et calcul differentiel exterieur. 323 565-570, C.R.A.S. Paris.

Department of Employment, Statistics Division, H. M. S. O. (1968).

Grandmont, J. M. (1992). Transformation of the commodity space, behavioral heterogeneity, and the aggregation problem. *Journal of Economic Theory*, 57, 1–35.

Hildenbrand, W. (1983). On the Law of Demand. *Econometrica*, 5, 997–1019.

Härdle, W., H. W. and Jerison, M. (1991). Empirical evidence on the law of demand. *Econometrica*, 59, 1525–1549.

Jerison, M. (1999). Dispersed excess demands, the weak axim and uniqueness of equilibrium. *Journal of Mathematical Economics*. Special Issue on Aggregation, forthcoming.

Kneip, A. (1998). Behavioral heterogeneity and structural properties of aggregate demand. *Journal of Mathematical Economics*. Special Issue on Aggregation, forthcoming.

Mitjuschin, L. G. and Polterovich, W. M. (1978). Criteria for monotonicity of demand functions. *Ekonomika i Matematicheski Metody*, 14, 122–128.

Quah, J. (1991). The law of demand when income is price dependent. *Econometrica*, 65, 1421–1442.

Sonnenschein, H. (1973). Do Walrast' identity and continuity characterize the class of community excess demand functions? *Journal of Economic Theory*, 6, 345–354.