March 2004

Compensated Variation and Hicksian Choice Probabilities in Random Utility Models that are Nonlinear in Income

by

John K. Dagsvik

and

Anders Karlström

Abstract:

In this paper we discuss Hicksian demand and compensating variation in the context of discrete choice. We first derive Hicksian choice probabilities and the distribution of the (random) expenditure function in the general case when the utilities are nonlinear in income. We subsequently derive exact and simple formulae for the expenditure and choice probabilities under price (policy) changes conditional on the initial utility level. This is of particular interest for welfare measurement because it enables the researcher to compute the distribution of Compensating variation in a simple way. We also derive formulae for the joint distribution of expenditure, the choice before and after a policy change has been introduced.

Keywords: Random expenditure function, Compensated choice probabilities, Compensating variation. Equivalent variation.

JEL classification: C25, D61

Acknowledgement: We wish to thank Mårten Palme, Thor O. Thoresen and Steinar Strøm for useful discussions and comments that have improved the paper. We are particularly grateful to a referee whose suggestions have improved the exposition of the paper substantially. Thanks to Anne Skoglund for excellent word processing and proof reading.

Address: Corresponding author; John K. Dagsvik, Research Department, Statistics Norway, P.O.Box 8131 Dep., N-0033 Oslo, Norway. E-mail: john.dagsvik@ssb.no

1. Introduction

In this paper we discuss the properties of the expenditure function and compensated (Hicksian) demand within the theory of discrete choice with particular reference to the analysis of compensated variation.[1] Since the theory of discrete choice is based on a random utility formulation, it follows that the corresponding expenditure and demand functions are random. In this paper we follow the standard random utility framework as set out in, e.g., McFadden (1981). Specifically, we obtain explicit formulae for the Hicksian choice probabilities and the distribution of the expenditure function. Hicksian choice probabilities have been discussed also by Small and Rosen (1981), within a somewhat different framework. Subsequently, we derive the distribution of the expenditure and the demand under price- or policy changes conditional on the utility being equal to the initial utility level. An immediate implication is that we readily obtain exact and convenient formulae for the distribution of Equivalent Variation and Compensated Variation (cv), which facilitates the analysis of welfare effects of changing prices or other attributes associated with the choice alternatives.

When the (random) utility function is nonlinear in income analytic formulae for the distribution of cv has so far not been available. Different approximations have been suggested in the literature. Morey et al. (1993) have used a representative individual approximation, while McFadden (1999) has developed a Monte Carlo simulator for computing cv in random utility models which converges to the true distribution of cv. Using this simulation method, Herriges and Kling (1999) have investigated the empirical consequences of nonlinear income effects based on a particular empirical application. Aaberge et al. (1995) have used a Monte Carlo simulation to compute equivalent variation. For a state-of-the-art review, see Herriges and Kling (1999) or Karlström (2001).

In the case where the model belongs to the additive Generalized Extreme Value (GEV) class the formulae become particularly simple.[2] As is well known, the choice probabilities for the GEV model can be expressed by simple closed form formulae and when the (indirect) utility functions are linear in income, the so-called log-sum formula can be applied to calculate sound welfare measures. However, when the utility function is nonlinear in income, no analogue to the log-sum approach exists.

Let us briefly review the notion of cv in random utility models for discrete choice. Let U_j denote the utility of alternative j and assume that

$$U_j = v_j\left(w_j, y\right) + \varepsilon_j,$$

where y represents income, w_j is a vector of attributes including price associated with alternative j, $v_j(\cdot)$, $j = 1, 2, ...,$ are deterministic functions and ε_j, $j = 1, 2, ...,$ are random terms with joint distribution that does not depend on the structural terms $\{v_j(\cdot)\}$.[3] Then (ignoring the choice set in the notation) cv is defined implicitly through

$$\max_j \left(v_j \left(w_j^0, y^0 \right) + \varepsilon_j \right) = \max_j \left(v_j \left(w_j, y^1 - cv \right) + \varepsilon_j \right)$$

where $\{w_j^0, y^0\}$ represent initial attributes and income, and $\{w_j, y^1\}$ are the attributes and income implied by the policy. From an analytic point of view the difficulty of deriving a formula for the distribution of cv stems from the fact that when the new attributes $\{w_j\}$ are introduced then the alternative that yields maximum utility may be different from the one that maximized utility initially. In other words, the individual agent may switch from the alternative chosen initially to a new one, when the policy is introduced. In this paper it is assumed that the random terms $\{\varepsilon_j\}$ are not affected by the policy intervention. This seems reasonable if the error terms $\{\varepsilon_k\}$ characterize tastes. It is less reasonable if $\{\varepsilon_k\}$ also includes unmeasured attributes of alternatives, which may be altered by policy. In general, the initial error terms $\{\varepsilon_k^0\}$ (say), may thus differ from the error terms $\{\varepsilon_k\}$ (say), after policy. The dependence between $\{\varepsilon_k^0\}$ and $\{\varepsilon_k\}$ will of course depend on the interpretation and modelling assumptions. For more discussion on this issue, see Heckman and Honoré (1990), McFadden (1999) and Carneiro et al. (2001).

One particular interpretation not explicitly covered by the above mentioned authors regards the evaluation of cv when some time has elapsed after the policy were introduced. In this case tastes may have changed from their initial values due to psychological factors. Dagsvik (2002) has considered discrete choice behavior in this type of setting. He deduces an explicit representation of the dependence between the error terms in two points in (continuous) time from an intertemporal version of the IIA assumption. This extended version of IIA accommodates serially dependent error terms due to what can be interpreted as taste persistence.

In addition to analysing the properties of cv, we derive formulae for the joint distribution of cv, the initial choice, and the choice after the policy intervention. This is useful in situations where one wishes to analyze how a specific policy may induce transitions from the initial chosen alternative to a new alternative, given that the utility level is kept unchanged. The distribution of cv is also useful for computing welfare measures based on weighted population means, as discussed by Hammond (1990).

When the utility function is linear (or separable) in income then the calculation of the mean cv becomes simple, at least within the GEV class, as mentioned above.

The paper is organized as follows. In Section 2 the discrete choice framework is presented, and in Section 3 compensating choice probabilities and the random expenditure function are defined and the corresponding distribution functions are derived. In Section 4 we derive compensated choice probabilities and the distribution of the expenditure function under price changes conditional on a utility level equal to the initial level under different assumptions about the random terms of the utility function. In Section 5 we discuss how the results obtained above can be extended to the case where the structural part $\left(v_j\right)$ of the utility function depends on random coefficients. In Section 6 we treat the special case where the random terms of the utility function are multivariate extreme value distributed, and in Section 7 we consider the case with independent and identically extreme value distributed random terms. Section 8 discusses the application of the results obtained to particular examples.

2. The setting

We consider a setting in which a consumer faces a set B of feasible alternatives (products), which is a subset of the universal set S of alternatives, $S=\{1,2,...,M\}$. The consumer's utility function of alternative j is assumed to have the form

(1) $$U_j = v_j\left(w_j, y\right) + \varepsilon_j$$

where $y > 0$ denotes income and w_j is a vector of attributes including price of alternative j. The function $v_j(\cdot)$ is assumed to be continuous, decreasing in the first argument and strictly increasing in the second, and it may depend on j.

Let $F^B(\cdot)$ denote the joint cumulative distribution function of $\{\varepsilon_k, k \in B\}$. We assume that $F(\cdot) \equiv F^S(\cdot)$ possesses a continuous density. Thus the probability of ties is zero.

We shall first assume that the joint distribution of $\{\varepsilon_k\}$ does not depend on $\{v_k\left(w_k, y\right)\}$. This may, however, be restrictive in some applications. Suppose for example that $v_j\left(w_j, y\right) = v_j\left(w_j, y; \beta\right)$ where β is a random coefficient. In this case the utility structure can be expressed as

$$U_j = E_\beta v_j\left(w_j, y; \beta\right) + \varepsilon_j^*$$

where E_β denotes the expectation with respect to β and the new error term ε_j^* is defined by

$$\varepsilon_j^* = v_j\left(w_j, y; \beta\right) - E_\beta v_j\left(w_j, y; \beta\right) + \varepsilon_j.$$

We thus realize that in this case the error terms $\left\{\varepsilon_k^*\right\}$ become dependent on $\left\{v_k\left(w_k, y;\beta\right)\right\}$. In Section 5 we discuss how to treat cases like this.

It is well known that one can express the Marshallian choice probabilities (i.e. the choice probabilities that correspond to Marshallian demand) by a simple formula. For notational simplicity, write $B = \left\{1, 2, ..., m\right\}$. Then the Marshallian choice probabilities are given by

(2) $$P_B\left(j, \mathbf{w}, y\right) \equiv P\left(U_j = \max_{k \in B} U_k\right) = \int F_j^B\left(u - v_1\left(w_1, y\right), u - v_2\left(w_2, y\right), ..., u - v_m\left(w_m, y\right)\right) du$$

where $F_j^B\left(x_1, x_2, ..., x_m\right)$ denotes the partial derivative with respect to x_j and $\mathbf{w} = \left(w_1, w_2, ..., w_m\right)$. Here it is understood that income and prices, (y,**w**) are *given*.

If $\left\{\varepsilon_k, k \in B\right\}$ are random variables with multivariate extreme value distribution, then G^B defined by

(3) $$\exp\left(-G^B\left(x_1, x_2, ..., x_m\right)\right) \equiv P\left(\bigcap_{k \in B}\left(\varepsilon_k \le x_k\right) \middle| y, \mathbf{w}\right)$$

has the property[4]

(4) $$G^B\left(x_1, x_2, ..., x_m\right) = e^{-z} G^B\left(x_1 - z, x_2 - z, ..., x_m - z\right)$$

for $z \in R$. The corresponding Marshallian choice probability is given by

(5) $$P_B\left(j, \mathbf{w}, y\right) = -\frac{G_j^B\left(-v_1\left(w_1, y\right), -v_2\left(w_2, y\right), ..., -v_m\left(w_m, y\right)\right)}{G^B\left(-v_1\left(w_1, y\right), -v_2\left(w_2, y\right), ..., -v_m\left(w_m, y\right)\right)}.$$

This formula is well known and is found in a completely equivalent form in for example McFadden (1981).

Remark

When the joint distribution of $\left(U_1, U_2, ..., U_M\right)$ is a multivariate extreme value distribution, type III, then the marginal distribution of U_j has the form

$$P\left(U_j \le u\right) = \exp\left(-\exp\left(v_j - u\tau\right)\right)$$

for $u \in R$, where v_j and $\tau > 0$ are parameters.[5] This implies that one can write

$$\tau U_j = v_j + \varepsilon_j$$

where

$$P\left(\varepsilon_j \le x\right) = \exp\left(-e^{-x}\right).$$

This means that when $\left(U_1, U_2, ..., U_M\right)$ is multivariate extreme value distributed the utility function has an *additive* structure in the sense that each random term is distributed according to the standardized (type III) extreme value distribution $\exp\left(-e^{-x}\right)$.

3. The random expenditure function and Hicksian (compensated) choice probabilities

We now proceed to discuss the notion of random expenditure function that corresponds to the above setting.

It is immediate from (1) that the indirect utility function

(6) $$V_B\left(\mathbf{w}, y\right) \equiv \max_{k \in B}\left(v_k\left(w_k, y\right) + \varepsilon_k\right)$$

has (for given $\left\{\varepsilon_k\right\}$) the standard properties of such functions deduced by Roy (1947), and exploited by McKenzie (1957), Diewert (1974), Varian (1992) and others. Specifically, so long as local non-satiation holds, $V_B\left(\mathbf{w}, y\right)$ is quasi-convex and homogenous of degree zero in prices and income, non-increasing in prices, and increasing in income, and can be inverted to give an expenditure function $Y_B\left(\mathbf{w}, u\right)$ satisfying the identity

(7) $$u = V_B\left(\mathbf{w}, Y_B\left(\mathbf{w}, u\right)\right).$$

The function $Y_B\left(\mathbf{w}, u\right)$ is concave and linear homogeneous in prices and increasing in u, and by Shephard's lemma satisfies almost surely the property that its price derivatives exist almost everywhere and equal the Hicksian demands (provided suitable differentiability conditions hold). These propositions hold under very general conditions, including the case of discrete alternatives; see Diamond and McFadden (1974), and McFadden (1978b).

The general properties listed above are, however, not immediately practical for deriving the *distributional* properties of the expenditure function and the Hicksian choice probabilities. Instead of starting by defining the expenditure function through (7) we find it more convenient and intuitive to start with the slightly more rigorous formulation

(8) $$Y_B(\mathbf{w},u)=\{z: V_B(\mathbf{w},z)=u\}.$$

In (8) the expenditure function is given as a *set*, but we shall see below that this set is a singleton. The expenditure function given in (8) can be readily computed as follows. Let $Y_k(w_k,u-\varepsilon_k)$ be determined by

(9) $$v_k(w_k,Y_k(w_k,u-\varepsilon_k))+\varepsilon_k=u.$$

Due to the fact that $v_k(w_k,y)$ is strictly increasing in y, $Y_k(w_k,u-\varepsilon_k)$ is uniquely determined. The interpretation of $Y_k(w_k,u-\varepsilon_k)$ is as the expenditure required to achieve utility level u, *given* alternative k with attributes w_k. We realize now that the expenditure function can be expressed as

(10) $$Y_B(\mathbf{w},u)=\min_{k\in B} Y_k(w_k,u-\varepsilon_k).$$

It therefore follows that with probability one the set $Y_B(\mathbf{w},u)$ is a *singleton.* Thus, the expenditure function can be defined uniquely by (8).

We shall see below that the setup above is very useful.

Theorem 1

*Let $B=\{1,2,...,m\}$. The expenditure function, $Y_B(\mathbf{w},u)$, is uniquely defined by (8), continuous in $(\mathbf{w},u)$, increasing in the price components of **w** and strictly increasing in u. Moreover, if $v_k(w_k,y)$ is concave in the price component for all k, $Y_B(\mathbf{w},u)$ is concave in prices. For $y_1>0,y_2>0,...,y_m>0$,*

(11) $$\begin{aligned}&P\big(Y_1(w_1,u-\varepsilon_1)>y_1,Y_2(w_2,u-\varepsilon_2)>y_2,...,Y_m(w_m,u-\varepsilon_m)>y_m\big)\\&=F^B\big(u-v_1(w_1,y_1),u-v_2(w_2,y_2),...,u-v_m(w_m,y_m)\big).\end{aligned}$$

Furthermore, the distribution of the expenditure function is given by

(12) $$P(Y_B(\mathbf{w},u)\le y)=1-F^B\big(u-v_1(w_1,y),u-v_2(w_2,y),...,u-v_m(w_m,y)\big),$$

for $u\in R, y>0$.

For the sake of bringing out the central arguments we have chosen to present the proof below instead of deferring it to the appendix.

Proof:

From (9) it follows that since $v_k(w_k, y)$ is continuous in y and the price component of w_k, decreasing in the price component of w_k and strictly increasing in y that $Y_k(w_k, x)$ is continuous in x and the price component of w_k, increasing in the price component of w_k and strictly increasing in x; see for example Rudin (1976), Theorem 4.17. Hence $Y_k(w_k, u-\varepsilon_k)$ is continuous in u and the price component of w_k, increasing in the price component of w_k and strictly increasing in u. It follows immediately from duality theory that if $v_k(w_k, y)$ is convex in the price component then $Y_k(w_k, x)$ will be concave in the price component. Due to the fact that B is finite it follows from (10) that $Y_B(\mathbf{w}, u)$ will also be continuous in u and the price components of **w**, increasing in the price components of **w** and strictly increasing in u. Since minimum of concave functions is concave (10) also implies that $Y_B(\mathbf{w}, u)$ is concave in prices. Furthermore, since $v_k(w_k, y)$ is strictly increasing in y we realize that (9) implies that

$$\left\{Y_k(w_k, u-\varepsilon_k) > y\right\} \Leftrightarrow \left\{v_k(w_k, y) + \varepsilon_k < u\right\}.$$

Hence, for $y_1 > 0, y_2 > 0, ..., y_m > 0$, we get from (10) that

$$P\left(\bigcap_{k\in B}\left(Y_k(w_k, u-\varepsilon_k) > y_k\right)\right) = P\left(\bigcap_{k\in B}\left(v_k(w_k, y_k) + \varepsilon_k < u\right)\right)$$

$$= F^B\left(u - v_1(w_1, y_1), u - v_2(w_2, y_2), ..., u - v_m(w_m, y_m)\right),$$

which proves (11). Eq. (12) then follows immediately from (10) and (11) by setting $y_k = y$ for all k.

Q.E.D.

The result in (12) is quite intuitive since (9) and (10) yield that

$$P\left(Y_B(\mathbf{w}, u) \le y\right) = P\left(V_B(\mathbf{w}, y) \ge u\right).$$

The event $\left\{Y_B(\mathbf{w},u) \le y\right\}$ means that the amount y is higher than or equal to the expenditure required to achieve utility level u. Evidently, this event is equivalent to $\left\{V_B(\mathbf{w},y) \ge u\right\}$. The latter event is the statement that the utility implied by income y is higher than or equal to u.

By imposing differentiability conditions on $\left\{v_k(w_k,y)\right\}$ one can sharpen the result of Theorem 1 to yield almost surely differentiable expenditure function in prices, as mentioned above.

The notion of Hicksian- or compensated choice probabilities does not seem to have appeared previously in the literature. Our definition of this concept follows next. Let $J_B(\mathbf{w},y)$ denote the choice from B given attributes and income $(\mathbf{w},y)$.

Definition 1

By Hicksian choice probabilities, $\left\{P_B^h(j,\mathbf{w},u)\right\}$*, we mean*

$$P_B^h(j,\mathbf{w},u) \equiv P\left(J_B\left(\mathbf{w},Y_B(\mathbf{w},u)\right) = j\right).$$

The interpretation of $P_B^h(j,\mathbf{w},u)$ is as the probability of choosing $j \in B$ given that the utility level is given and equal to u. For example, if prices change the consumers are ensured income compensation so as to maintain a given utility level.

Note that the Hicksian choice probabilities can also be expressed as

$$\text{(13)} \qquad P_B^h(j,\mathbf{w},u) = P\left(Y_j\left(w_j,u-\varepsilon_j\right) = Y_B(\mathbf{w},u)\right) = P\left(Y_j\left(w_j,u-\varepsilon_j\right) = \min_{k\in B} Y_k\left(w_k,u-\varepsilon_k\right)\right).$$

Theorem 2

The Hicksian choice probabilities can be expressed as

$$P_B^h(j,\mathbf{w},u) = \int_0^\infty F_j^B\left(u-v_1(w_1,y),u-v_2(w_2,y),\ldots,u-v_m(w_m,y)\right)v_j\left(w_j,dy\right)$$

for $u \in R$.

Proof:

From (13) and (11) in Theorem 1 we obtain that

$$
\begin{aligned}
&P\left(Y_j\left(w_j,u-\varepsilon_j\right)=Y_B\left(\mathbf{w},u\right),Y_B\left(\mathbf{w},u\right)\in\left(y,y-dy\right)\right)\\
&=P\left(Y_j\left(w_j,u-\varepsilon_j\right)\in\left(y,y-dy\right),\min_{k\in B\setminus\{j\}}Y_k\left(w_k,u-\varepsilon_k\right)\geq y\right)\\
&=P\left(Y_j\left(w_j,u-\varepsilon_j\right)\in\left(y-dy,y\right),\bigcap_{k\in B\setminus\{j\}}\left(Y_k\left(w_k,u-\varepsilon_k\right)\geq y\right)\right)\\
&=F_j^B\left(u-v_1\left(w_1,y\right),u-v_2\left(w_2,y\right),...,u-v_m\left(w_m,y\right)\right)v_j\left(w_j,dy\right).
\end{aligned}
\tag{14}
$$

The result now follows by integration with respect to y.

Q.E.D.

From Theorem 2 we realize that one can calculate the Hicksian choice probabilities readily provided the cumulative distribution $F^B(\cdot)$ is known since only a one dimensional integral is involved in the formulae for $P_B^h\left(j,\mathbf{w},u\right)$.

To bring out the symmetry of the Marshallian and Hicksian choice probabilities, recall that similarly to (14) one has that

$$P\left(J_B\left(\mathbf{w},y\right)=j,V_B\left(\mathbf{w},y\right)\in\left(u,u+du\right)\right)=F_j^B\left(u-v_1\left(w_1,y\right),u-v_2\left(w_2,y\right),...,u-v_m\left(w_m,y\right)\right)du\ , \tag{15}$$

from which the corresponding Marshallian choice probability follows by integration. Thus the only difference between (14) and (15) is the "Jacobian" , $v_j\left(w_j,dy\right)$, associated with the choice j. This Jacobian is due to the change of variable from u_j to y_j, where $u_j=v_j\left(w_j,y_j\right)$.

The next result is useful for calculating moments of the expenditure function.

Lemma 1

Let H be a probability distribution. Then for any $\alpha\geq 1$

$$\int_0^\infty x^\alpha d\,H\left(x\right)=\alpha\int_0^\infty x^{\alpha-1}\left(1-H\left(x\right)\right)dx$$

where the two sides exists or diverge together.

The result of Lemma 1 is well known, but for the reader's convenience we provide a proof in the appendix.

Corollary 1 (Shephard's Lemma)

Suppose $v_j\left(w_j, y\right)=\psi_j\left(y-w_{1j}, w_{2j}\right)$*, for some suitable function* ψ_j *that is continuously differentiable in the first argument, where* w_{1j} *is the price (user cost) associated with alternative j and* w_{2j} *denotes other attributes. Suppose furthermore that* $EY_k\left(w_k, u-\varepsilon_k\right)<\infty$ *for all* $k\in S$*. Then* $EY_B\left(\boldsymbol{w}, u\right)<\infty$*, and*

$$\frac{\partial EY_B\left(\boldsymbol{w},u\right)}{\partial w_{1j}}=P_B^h\left(j,\boldsymbol{w},u\right).$$

Proof:

Note first that

$$\frac{\partial v_j\left(w_j,y\right)}{\partial w_{1j}}=-\psi_{j1}\left(y-w_{1j},w_{2j}\right)=\frac{-\partial v_j\left(w_j,y\right)}{\partial y}$$

where $\psi_{j1}\left(x,w_{2j}\right)$ denotes the derivative with respect to $x, x\in[0,\infty)$. Since $EY_B\left(\mathbf{w},u\right)\le EY_k\left(w_k,u-\varepsilon_k\right)$ for any $k\in B$ the expectation of the expenditure function is finite. Hence, by Lemma 1 and (12)

$$\frac{\partial EY_B\left(\mathbf{w},u\right)}{\partial w_{1j}}=\partial\left[\int_0^\infty F^B\left(u-v_1\left(w_1,y\right),u-v_2\left(w_2,y\right),...,u-v_m\left(w_m,y\right)\right)dy\right]\Big/\partial w_{1j}$$

$$=\int_0^\infty F_j^B\left(u-v_1\left(w_1,y\right),u-v_2\left(w_2,y\right),...,u-v_m\left(w_m,y\right)\right)\frac{\partial v_j\left(w_j,y\right)}{\partial y}dy=P_B^h\left(j,\mathbf{w},u\right).$$

Above, integration under the integral sign is possible, since the integrand is continuous and the integral evidently converges.

Q.E.D.

We recognize the result of Corollary 1 as a probabilistic (or aggregate) version of *Shephard's Lemma*. It states that the partial derivative of the aggregate expenditure function with respect to the price of alternative j yields the fractional compensated demand. The structure $\psi_j\left(y-w_{1j},w_{2j}\right)$ of the systematic part of the utility of alternative j is quite general and covers the standard discrete choice settings. This is in contrast to the corresponding duality result in the Marshallian case. As McFadden (1981) has discussed, the Marshallian choice probabilities follow from the mean indirect utility function, by means of Roy's identity, only when utility is linear in income.

4. The probability distribution of the expenditure function and the choice under price changes conditional on the initial utility level

We shall next consider the problem of characterizing the distribution of the expenditure function and the choice probabilities when the utility level equals the (indirect) utility under prices and income that differ from the current prices and incomes. To this end we consider a two period setting. In period one (the initial period) the attributes and income are $(\mathbf{w}^0, y^0)$. In the second period (current period) the attributes and income are $(\mathbf{w}, y^1)$. As above, it is assumed that the respective random terms remain unchanged under attribute changes. In general, when attributes change it may yield a decrease or an increase in the agent's indirect utility. Furthermore, the highest utility may no longer be attained at the alternative chosen initially, and consequently the agent will switch to a new alternative, namely the one that maximizes utility under the new attribute regime. In the current setting, however, the (indirect) utility level is kept fixed and equal to the initial level. But the agent may still switch from the initially chosen alternative to a new one because, after the attributes change, the utility of the initially chosen alternatives may no longer coincide with the new indirect utility.

Let us first consider the joint distribution of the initial choice and the current expenditure given that the utility level is equal to the initial utility level. Formally, this is the joint distribution of $\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right), J_B\left(\mathbf{w}^0, y^0\right)\right)$. The interpretation of $Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)$ is as the expenditure function conditional on the utility level that corresponds to income level y^0.

Theorem 3

Let $y_i\left(w_i^0, y^0, w_i\right)$ and $h_i\left(w_i^0, y^0, w_i, y\right)$ be defined by

$$v_i\left(w_i^0, y^0\right) = v_i\left(w_i, y_i\left(w_i^0, y^0, w_i\right)\right)$$

and

$$h_i\left(w_i^0, y^0, w_i, y\right) = max\left(v_i\left(w_i, y\right), v_i\left(w_i^0, y^0\right)\right).$$

The joint distribution of $Y_B\left(\boldsymbol{w}, V_B\left(\boldsymbol{w}^0, y\right)\right)$ and $J_B\left(\boldsymbol{w}^0, y^0\right)$ is given by

$$P\left(Y_B\left(\boldsymbol{w},V_B\left(\boldsymbol{w}^0,y^0\right)\right)>y,J_B\left(\boldsymbol{w}^0,y^0\right)=i\right)$$
$$=P\left(h_i\left(w_i^0,y^0,w_i,y\right)+\varepsilon_i=\max_{k\in B}\left(h_k\left(w_k^0,y^0,w_k,y\right)+\varepsilon_k\right)\right)$$
$$=\int F_i^B\left(u-h_1\left(w_1^0,y^0,w_1,y\right),u-h_2\left(w_2^0,y^0,w_2,y\right),\ldots,u-h_m\left(w_m^0,y^0,w_m,y\right)\right)du$$

for $i\in B$, *and* $0<y<y_i\left(w_i^0,y^0,w_i\right)$. *When* $y\geq y_i\left(w_1^0,y^0,w_i\right)$

$$P\left(Y_B\left(\boldsymbol{w},V_B\left(\boldsymbol{w}^0,y^0\right)\right)>y,\,J_B\left(\boldsymbol{w}^0,y^0\right)=i\right)=0\,.$$

From Theorem 3 we notice that the joint distribution of the expenditure, given the initial utility level, and the initial choice, can be expressed as a *choice probability*.

Although the result of Theorem 3 follows from Theorem 4 below we have given an independent proof of Theorem 3 in the appendix. This is of interest because it demonstrates that if one is only interested in the distribution of the expenditure function there is no need to proceed via the result of Theorem 4, which is more complicated to prove than the result of Theorem 3.

The intuition of the result of Theorem 3 can be perceived as follows: Since

$$\left\{Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)>y\right\}\Leftrightarrow\left\{V_B\left(\mathbf{w},y\right)<V_B\left(\mathbf{w}^0,y^0\right)\right\}$$

it follows that the alternative that is chosen initially has utility that is the highest utility initially and also higher than the highest utility in the second period. Hence, if alternative i is chosen initially and current expenditure is higher than y it must be true that $v_i\left(w_i^0,y^0\right)>v_i\left(w_i,y\right)$, and

$$v_i\left(w_i^0,y^0\right)+\varepsilon_i=h_i\left(w_i^0,y^0,w_i,y\right)+\varepsilon_i=\max\left(V_B\left(\mathbf{w},y\right),V_B\left(\mathbf{w}^0,y^0\right)\right)=\max_{k\in B}\left(h_k\left(w_k^0,y^0,w_k,y\right)+\varepsilon_k\right).$$

From Theorem 3 we immediately obtain the conditional distribution of $Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)$ given the initial choice $J_B\left(\mathbf{w}^0,y^0\right)$. Note also that $y<y_i\left(w_i^0,y^0,w_i\right)$ implies that

$$h_i\left(w_i^0,y^0,w_i,y\right)=v_i\left(w_i^0,y^0\right).$$

To state the next result it is convenient to apply the following notation

$$I_k\left(w_k^0,y^0,w_k,y\right)=\begin{cases}1 & \text{if}\quad v_k\left(w_k,y\right)<v_k\left(w_k^0,y^0\right),\\ 0 & \text{otherwise}.\end{cases}\tag{16}$$

The next corollary follows from Theorem 3 by summing over $i \in B$.

Corollary 2

The distribution of $Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)$ *is given by*

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right)$$
$$= \sum_{i \in B} I_i\left(w_i^0, y^0, w_i, y\right) \int F_i^B\left(u - h_1\left(w_1^0, y^0, w_1, y\right), u - h_2\left(w_2^0, y^0, w_2, y\right), \ldots, u - h_m\left(w_m^0, y^0, w_m, y\right)\right) du$$

for $y > 0$.

Note that it follows from Theorem 3 that

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right) = 0$$

when

$$y \geq \max_{i \in B} y_i\left(w_i^0, y^0, w_i\right).$$

Recall that the Compensation Variation measure (cv) is defined as

$$cv = y^1 - Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}, y^0\right)\right).$$

The distribution of cv thus follows directly from Corollary 2.

Let $J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right)$ denote the current choice from B, given the current and initial prices and income $\left(\mathbf{w}, \mathbf{w}^0, y^0\right)$, and given that the current utility level equals the initial one, $V_B\left(\mathbf{w}^0, y^0\right)$. Thus, $J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right)$ is defined by

$$J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right) = J_B\left(\mathbf{w}, Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)\right).$$

Let us next consider the joint distribution of the current expenditure, the current and the initial choice, given that the utility level is kept equal to the initial utility level. That is, we shall consider the joint distribution of $\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right), J_B\left(\mathbf{w}^0, y^0\right), J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right)\right)$.

The reason why we consider the joint distribution of the expenditure and the initial and current choices, given that the utility level is equal to the initial utility level, is that it may be of interest in policy simulations because it enables us to calculate cv conditional on the initial- or current choice, or both.

Theorem 4

Let

$$C\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}, y\right) = \left\{k : \mathrm{v}_k\left(w_k, y\right) \geq \mathrm{v}_k\left(w_k^0, y^0\right)\right\}.$$

We have that

$$P\left(Y_B\left(\boldsymbol{w}, V_B\left(\boldsymbol{w}^0, y^0\right)\right) \in dy, J_B\left(\boldsymbol{w}^0, y^0\right) = i, J_B^*\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}\right) = j\right)$$
$$= \mathrm{v}_j\left(w_j, dy\right) \int F_{ij}^B\left(u - h_1\left(w_1^0, y^0, w_1, y\right), u - h_2\left(w_2^0, y^0, w_2, y\right), \ldots, u - h_m\left(w_m^0, y^0, w_m, y\right)\right) du$$

when $i \in B \setminus C\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}, y\right)$, $j \in C\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}, y\right)$, *where* F_{ij} *denotes the derivative of F with respect to components i and j. When* $i \notin B \setminus C\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}, y\right)$, $j \notin C\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}, y\right)$

$$P\left(Y_B\left(\boldsymbol{w}, V_B\left(\boldsymbol{w}^0, y^0\right)\right) \in dy, J_B\left(\boldsymbol{w}^0, y^0\right) = i, J_B^*\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}\right) = j\right) = 0,$$

For $j = i \in B$, *and* $y = y_i\left(w_i^0, y^0, w_i\right)$,

$$P\left(Y_B\left(\boldsymbol{w}, V_B\left(\boldsymbol{w}^0, y^0\right)\right) = y, J_B\left(\boldsymbol{w}^0, y^0\right) = J_B^*\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}\right) = i\right) = P_B\left(i, \boldsymbol{w}, y\right).$$

When $y \neq y_i\left(w_i^0, y^0, w_i\right)$,

$$P\left(Y_B\left(\boldsymbol{w}, V_B\left(\boldsymbol{w}^0, y^0\right)\right) \in dy, J_B\left(\boldsymbol{w}^0, y^0\right) = J_B^*\left(\boldsymbol{w}^0, y^0, \boldsymbol{w}\right) = i\right) = 0.$$

A proof of Theorem 4 is given in the appendix.

Note that we have expressed the result of Theorem 4 as a differential of the c.d.f. If $\mathrm{v}_j\left(\mathrm{w}_j, \mathrm{y}\right)$ is differentiable with respect to income the corresponding joint density exists.

When y is such that $C\left(\mathbf{w}^0, y^0, \mathbf{w}, y\right)$ and $B \setminus C\left(\mathbf{w}^0, y^0, \mathbf{w}, y\right)$ are non-empty, the agent must switch from the initial chosen alternative i (say) in $B \setminus C\left(\mathbf{w}^0, y^0, \mathbf{w}, y\right)$ to an alternative within $C\left(\mathbf{w}^0, y^0, \mathbf{w}, y\right)$ to achieve $Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) \in dy$, unless $y = y_i\left(w_i^0, y^0, w_i\right)$. This is so because $B \setminus C\left(\mathbf{w}^0, y^0, \mathbf{w}, y\right)$ contains the most attractive alternatives in the initial period while $C\left(\mathbf{w}^0, y^0, \mathbf{w}, y\right)$ contains the most attractive alternatives in the current period. Equivalently, if $j \neq i$ then

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) \in dy, J_B\left(\mathbf{w}^0, y^0\right) = i, J_B^*\left(\mathbf{w}, y^0, \mathbf{w}\right) = j\right) > 0$$

if $y_j\left(w_j^0, y^0, w_j\right) < y_i\left(w_i^0, y^0, w_i\right)$ and $y_j\left(w_j^0, y^0, w_j\right) \leq y < y_i\left(w_i^0, y^0, w_i\right)$. Otherwise this probability equals zero.

The result of Theorem 3 shows that only a one-dimensional integral is needed to calculate the joint probability density of

$$\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right), J_B\left(\mathbf{w}^0, y^0\right), J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right)\right)$$

provided $F_{ij}^B\left(x_1, ..., x_m\right)$ is known. However, in cases where one only has closed form expressions for the density of $F^B\left(x_1, x_2, ..., x_m\right)$, such as in the Multinomial Probit case, a $m-2$ dimensional integral is needed to calculate $F_{ij}^B\left(x_1, ..., x_m\right)$.

The result of Theorem 3 enables us to calculate compensating "transition" probabilities given by

$$P\left(J_B\left(\mathbf{w}^0, y^0\right) = i, J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right) = j\right).$$

This expression represents the probability of going from i to j when attributes change from $\mathbf{w}^0$ to $\mathbf{w}$, given that the utility level is kept fixed and equal to the initial utility, $V_B\left(\mathbf{w}^0, y^0\right)$, and given that the error terms remain unchanged. Specifically, we have that the joint probability of choosing i initially and j in the current period equals

$$P\left(J_B\left(\mathbf{w}^0, y^0\right) = i, J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right) = j\right) = \int_{y_j\left(w_j^0, y^0, w_j\right)}^{y_i\left(w_i^0, y^0, w_i\right)} P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) \in dy, J_B\left(\mathbf{w}^0, y^0\right) = i, J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right)\right)$$

where the integrand in the above integral is given in Theorem 4. From this expression the conditional probabilities of the current choice given the initial choice follows. In some cases it may be of interest to calculate the fractions of such “compensated” transitions that follow from a specific policy intervention. Furthermore, one can similarly calculate the change in the choice probability given that the utility level is kept equal to the initial utility level. The latter expression equals

$$P\left(J_B^*\left(\mathbf{w}^0, y^0, \mathbf{w}\right) = j\right) - P\left(J_B\left(\mathbf{w}^0, y^0\right) = j\right)$$

for $j \in B$.

Remark

The results obtained in Theorem 3 and Theorem 4 are derived under the assumption that the choice set B is the same before and after the price change. However, these results can be slightly modified to apply also in cases where the choice set changes. Suppose for example that alternative 2 was available initially but is removed as part of a policy intervention. One can conveniently accommodate for this by letting w_2 become very large so that $v_2\left(w_2, y\right)$ becomes very small. As a result we obtain that

$$h_2\left(w_2^0, y, w_2, y\right) = v_2\left(w_2^0, y^0\right)$$

and that

$$2 \notin C\left(y^0, y, \mathbf{w}^0, \mathbf{w}\right).$$

From Theorem 3 and Lemma 1 it follows that the mean and variance of $Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)$ can be calculated by the formulae

$$\text{(17)} \quad \begin{aligned} & E\,Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) \\ & = \sum_{i \in B} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \int F_i^B\left(u - h_1\left(w_1^0, y^0, w_1, y\right), u - h_2\left(w_2^0, y^0, w_2, y\right), \ldots, u - h_m\left(w_m^0, y^0, w_m, y\right)\right) du\, dy, \end{aligned}$$

and

$$E\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)\right)^2$$

(18)

$$= 2\sum_{i\in B} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} y \int F_i^B\left(u - h_1\left(w_1^0, y^0, w_1, y\right), u - h_2\left(w_2^0, y^0, w_2, y\right), ..., u - h_m\left(w_m^0, y^0, w_m, y\right)\right) du\, dy.$$

5. Models with random coefficients

Above we assumed that the random terms of the utility function were independent of the respective structural terms. We shall now relax this assumption. Specifically, we now suppose that

$$U_j = v_j\left(w_j, y; \beta\right) + \varepsilon_j$$

where the notation above means that the systematic part $v_j\left(w_j, y; \beta\right)$ depends on a vector of parameters β which are random and distributed on a suitable space. We assume, however, that β is independent of $(\varepsilon_1, \varepsilon_2, ...)$. A special case of this type of models is the so-called Mixed Multinomial Logit Model (MNL). The mixed MNL model has recently become popular because it provides a very general random utility modeling framework that is convenient to apply in empirical applications, see McFadden and Train (2000).

We realize that Theorem 3 still holds when β is given, i.e.,

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y, J_B\left(\mathbf{w}^0, y^0\right) = i \middle| \beta\right)$$
$$= \int F_i^B\left(u - h_1\left(w_1^0, y^0, w_1, y; \beta\right), u - h_2\left(w_2^0, y^0, w_2, y; \beta\right), ..., u - h_m\left(w_m^0, y^0, w_m, y; \beta\right)\right) du$$

for $0 < y < y_i\left(w_i^0, y^0.w_i; \beta\right)$, where $y_i\left(w_i^0, y^0, w_i; \beta\right)$ is determined by

$$v_i\left(w_i^0, y^0; \beta\right) = v_i\left(w_i^0, y_{ii}\left(w_i^0, y^0, w_i; \beta\right); \beta\right)$$

and

$$h_j\left(w_j^0, y^0, w_j, y; \beta\right) = \max\left(v_j\left(w_j, y; \beta\right), v_j\left(w_j^0, y^0; \beta\right)\right).$$

Consequently, we get

$$P\left(Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)>y,J_B\left(\mathbf{w}^0,y^0\right)=i\right)=EP\left(Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)>y,J_B\left(\mathbf{w}^0,y^0\right)=i\middle|\beta\right)$$
$$=\int I_i\left(w_i^0,y^0,w_i,y;\beta\right)F_i^B\left(u-h_1\left(w_1^0,y^0,w_1,y;\beta\right),u-h_2\left(w_2^0,y^0,w_2,y;\beta\right),\dots,u-h_m\left(w_m^0,y^0,w_m,y;\beta\right)\right)f(\beta)du\,d\beta$$

where $f(\beta)$ is the probability density of β and

$$I_i\left(w_i^0,y^0,w_i,y;\beta\right)=\begin{cases}1 & \text{if}\quad v_i\left(w_i,y;\beta\right)<v_i\left(w_i^0,y^0;\beta\right)\\ 0 & \text{otherwise}.\end{cases}$$

Note that the interpretation given in Theorem 3 still holds in this case.

The corresponding results for Theorem 4 and the results given above are completely analogous.

6. Specialization to the case with multivariate extreme value distributed error terms

McFadden (1978a) introduced the GEV class of models that follows if F is a multivariate extreme value distribution function. The GEV class represents no essential restrictions on the class of random utility models, since Dagsvik (1994, 1995) and Joe (2001) have demonstrated that one can approximate any random utility model arbitrarily closely by GEV models. This class contains the well known Multinomial Logit and Nested Logit models which has proven to be very important in applied work.

We shall state how the general results obtained in Theorem 4 simplifies in the case where $F(\cdot)$ is a multivariate extreme value distribution. For simplicity, we only state the joint density of

$$\left(Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right),J_B\left(\mathbf{w}^0,y^0\right),J_B^*\left(\mathbf{w}^0,y^0,\mathbf{w}\right)\right)$$

for the case when

$$J_B\left(\mathbf{w}^0,y^0\right)\neq J_B^*\left(\mathbf{w}^0,y^0,\mathbf{w}\right).$$

Corollary 3

Suppose $F(\cdot)$ is a multivariate extreme value distribution. Then

$$P\left(Y_B\left(\boldsymbol{w},V_B\left(\boldsymbol{w}^0,y^0\right)\right)\in dy,J_B\left(\boldsymbol{w}^0,y^0\right)=i,J_B^*\left(\boldsymbol{w}^0,y^0,\boldsymbol{w}\right)=j\right)$$
$$=\frac{\left(\tilde{G}^B\tilde{G}_{ij}^B-\tilde{G}_i^B\tilde{G}_j^B\right)\mathrm{v}_j\left(w_j,dy\right)}{\left(\tilde{G}^B\right)^2}$$

when $i\neq j, i,j\in B$ *and* $y_j\left(w_j^0,y^0,w_j\right)\leq y<y_i\left(w_i^0,y^0,w_i\right)$, *where*

$$\tilde{G}^B=G^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\ldots,-h_m\left(w_m^0,y^0,w_m,y\right)\right),$$

$$\tilde{G}_i^B=G_i^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\ldots,-h_m\left(w_m^0,y^0,w_m,y\right)\right),$$

$$\tilde{G}_{ij}^B=G_{ij}^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\ldots,-h_m\left(w_m^0,y^0,w_m,y\right)\right),$$

and $G^B(\cdot)$ *is defined in (3).*

The proof of Corollary 3 is given in the appendix.

Thus, the result of Corollary 3 implies that in the case where the random terms are multivariate extreme value distributed, one can rather easily compute the joint density of the expenditure function and J_B and J_B^*. Note that no integration is needed here.

The next corollary follows directly from Theorem 3.

Corollary 4

Suppose $F(\cdot)$ *is a multivariate extreme value distribution. Then we have, for* $i\in B$, *and* $0<y<y_i\left(w_i^0,y^0,w_i\right)$, *that*

$$P\left(Y_B\left(\boldsymbol{w},V_B\left(\boldsymbol{w}^0,y^0\right)\right)>y,J_B\left(\boldsymbol{w}^0,y^0\right)=i\right)$$
$$=-\frac{G_i^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\ldots,-h_m\left(w_m^0,y^0,w_m,y\right)\right)}{G^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\ldots,-h_m\left(w_m^0,y^0,w_m,y\right)\right)}$$

and

$$P\left(Y_B\left(\boldsymbol{w},V_B\left(\boldsymbol{w}^0,y^0\right)\right)>y\right)$$
$$=-\frac{\sum_{i\in B} I_i\left(w_i^0,y^0,w_i,y\right)G_i^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\dots,-h_m\left(w_m^0,y^0,w_m,y\right)\right)}{G^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\dots,-h_m\left(w_m^0,y^0,w_m,y\right)\right)}$$

for $y>0$.

From eq. (17), Lemma 1 and Corollary 4 we obtain that

(19)
$$\mathrm{E}Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)$$
$$=-\sum_{i\in B}\int_0^{y_i\left(w_i^0,y^0,w_i\right)}\frac{G_i^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\dots,-h_m\left(w_m^0,y^0,w_m,y\right)\right)dy}{G^B\left(-h_1\left(w_1^0,y^0,w_1,y\right),-h_2\left(w_2^0,y^0,w_2,y\right),\dots,-h_m\left(w_m^0,y^0,w_m,y\right)\right)}.$$

Thus, formulae (19) can be applied to compute the mean Compensating Variation

$$y^1-\mathrm{E}\,Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)$$

under the assumption of Corollary 4.

7. Specialization to the case with i.i. extreme value distributed error terms

When the error terms are i.i. extreme value distributed the Marshallian choice probabilities reduce to the Luce model- or the Multinomial Logit model. In this case the distribution of $Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)$ becomes particularly simple.

From Corollary 4 we immediately get the next result.

Corollary 5

Suppose $\varepsilon_1,\varepsilon_2,\dots,\varepsilon_M$ *, are independent and standard extreme value distributed. Then for* $i\in B$ *, and* $0<y<y_i\left(w_i^0,y^0,w_i\right)$

$$P\left(Y_B\left(\boldsymbol{w},V_B\left(\boldsymbol{w}^0,y^0\right)\right)>y,J_B\left(\boldsymbol{w}^0,y^0\right)=i\right)=\frac{exp\left(\mathrm{v}_i\left(w_i^0,y^0\right)\right)}{\sum_{k\in B} exp\left(max\left(\mathrm{v}_k\left(w_k,y\right),\mathrm{v}_k\left(w_k^0,y^0\right)\right)\right)}$$

and

$$P\left(Y_B\left(\boldsymbol{w}, V_B\left(\boldsymbol{w}^0, y^0\right)\right) > y\right) = \frac{\sum_{i\in B} I_i\left(w_i^0, y^0, w_i, y\right) exp\left(\mathrm{v}_i\left(w_i^0, y^0\right)\right)}{\sum_{k\in B} exp\left(max\left(\mathrm{v}_k\left(w_k, y\right), \mathrm{v}_k\left(w_k^0, y^0\right)\right)\right)}$$

for $y > 0$.

From eq. (17), Lemma 1 and Corollary 5 we obtain that

$$\text{(20)}\quad \mathrm{E}\,\mathrm{Y}_\mathrm{B}\left(\mathbf{w}, \mathrm{V}_\mathrm{B}\left(\mathbf{w}^0, \mathrm{y}^0\right)\right) = \sum_{\mathrm{i}\in\mathrm{B}} \exp\left(\mathrm{v}_\mathrm{i}\left(\mathrm{w}_\mathrm{i}^0, \mathrm{y}^0\right)\right) \int_0^{\mathrm{y}_\mathrm{i}\left(\mathrm{w}_\mathrm{i}^0, \mathrm{y}^0, \mathrm{w}_\mathrm{i}\right)} \frac{\mathrm{dy}}{\sum_{\mathrm{k}\in\mathrm{B}} \exp\left(\max\left(\mathrm{v}_\mathrm{k}\left(\mathrm{w}_\mathrm{k}, \mathrm{y}\right), \mathrm{v}_\mathrm{k}\left(\mathrm{w}_\mathrm{k}^0, \mathrm{y}^0\right)\right)\right)}.$$

Similarly to eq. (19), eq. (20) can be applied to compute the mean Compensating Variation

$$\mathrm{y}^1 - \mathrm{E}\,\mathrm{Y}_\mathrm{B}\left(\mathbf{w}, \mathrm{V}_\mathrm{B}\left(\mathbf{w}^0, \mathrm{y}^0\right)\right)$$

in the case where the random utilities have i.i. extreme value distributed error terms.

8. Examples

Example 1

Consider a Nested Logit model with 4 alternatives where joint c.d.f. of the error terms of the utilities is given by

$$\text{(21)}\qquad \exp\left(-\mathrm{G}\left(\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3, \mathrm{x}_4\right)\right) = \exp\left(-\mathrm{e}^{-\mathrm{x}_1} - \mathrm{e}^{-\mathrm{x}_2} - \left(\mathrm{e}^{-\mathrm{x}_3/\theta} + \mathrm{e}^{-\mathrm{x}_4/\theta}\right)^\theta\right)$$

where $\theta \in (0,1]$ and $1-\theta^2$ has the interpretation as the correlation between the error terms of the utilities of alternatives three and four. For simplicity, write $\mathrm{h}_\mathrm{j}(\mathrm{y}) = \mathrm{h}_\mathrm{j}\left(\mathrm{w}_\mathrm{j}^0, \mathrm{y}^0, \mathrm{w}_\mathrm{j}, \mathrm{y}\right)$. We then get

$$-\frac{\mathrm{G}_\mathrm{i}\left(-\mathrm{h}_1(\mathrm{y}), -\mathrm{h}_2(\mathrm{y}), -\mathrm{h}_3(\mathrm{y}), -\mathrm{h}_4(\mathrm{y})\right)}{\mathrm{G}\left(-\mathrm{h}_1(\mathrm{y}), -\mathrm{h}_2(\mathrm{y}), -\mathrm{h}_3(\mathrm{y}), -\mathrm{h}_4(\mathrm{y})\right)} = \frac{\mathrm{e}^{\mathrm{h}_\mathrm{i}(\mathrm{y})}}{\mathrm{e}^{\mathrm{h}_1(\mathrm{y})} + \mathrm{e}^{\mathrm{h}_2(\mathrm{y})} + \left(\mathrm{e}^{\mathrm{h}_3(\mathrm{y})/\theta} + \mathrm{e}^{\mathrm{h}_4(\mathrm{y})/\theta}\right)^\theta}$$

for $\mathrm{i} \in \{1,2\}$, and

$$-\frac{\mathrm{G}_\mathrm{i}\left(-\mathrm{h}_1(\mathrm{y}), -\mathrm{h}_2(\mathrm{y}), -\mathrm{h}_3(\mathrm{y}), -\mathrm{h}_4(\mathrm{y})\right)}{\mathrm{G}\left(-\mathrm{h}_1(\mathrm{y}), -\mathrm{h}_2(\mathrm{y}), -\mathrm{h}_3(\mathrm{y}), -\mathrm{h}_4(\mathrm{y})\right)} = \frac{\left(\mathrm{e}^{\mathrm{h}_3(\mathrm{y})/\theta} + \mathrm{e}^{\mathrm{h}_4(\mathrm{y})/\theta}\right)^{\theta-1} \mathrm{e}^{\mathrm{h}_\mathrm{i}(\mathrm{y})/\theta}}{\mathrm{e}^{\mathrm{h}_1(\mathrm{y})} + \mathrm{e}^{\mathrm{h}_2(\mathrm{y})} + \left(\mathrm{e}^{\mathrm{h}_3(\mathrm{y})/\theta} + \mathrm{e}^{\mathrm{h}_4(\mathrm{y})/\theta}\right)^\theta}$$

for $i \in \{3,4\}$. Recall that $h_i(y) = v_i\left(w_i^0, y^0\right)$ when $I_i\left(w_i^0, y^0, w_i, y\right) = 1$. From Corollary 3 we get that

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right) \tag{22}$$
$$= \frac{\sum_{i=1}^{2} I_i\left(w_i^0, y^0, w_i, y\right) e^{v_i\left(w_i^0, y^0\right)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta}\right)^{\theta-1} \sum_{j=3}^{4} I_i\left(w_i^0, y^0, w_i, y\right) e^{v_i\left(w_i^0, y^0\right)/\theta}}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta}\right)^{\theta}},$$

for $y > 0$. Furthermore, the corresponding mean expenditure follows from (19) and equals

$$E\, Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) = \sum_{i=1}^{2} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \frac{e^{v_i\left(w_i^0, y^0\right)} dy}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta}\right)^{\theta}} \tag{23}$$
$$+ \sum_{i=3}^{4} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \frac{\left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta}\right)^{\theta-1} e^{v_i\left(w_i^0, y^0\right)/\theta} dy}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta}\right)^{\theta}}.$$

When $\theta = 1$ we get the Multinomial Logit model. In this case we immediately see that (22) and (23) are indeed consistent with Corollary 5 and (20).

Example 2

In this example we assume that a new alternative enters the choice set in the current period. The setting in the initial period is as in Example 1, while in the current period the joint c.d.f. is given by

$$\exp\left(-G\left(x_1, x_2, x_3, x_4, x_5\right)\right) = \exp\left(-\left(e^{-x_1} + e^{-x_2} + \left(e^{-x_3/\theta} + e^{-x_4/\theta} + e^{-x_5/\theta}\right)^{\theta}\right)\right) \tag{24}$$

where alternative 5 denotes the new alternative. Thus (21) follows from (24) by letting $x_5 = \infty$ in (24). As mentioned above, results obtained in this paper can still be used to find the distribution of the expenditure function. This is obtained formally by assuming that also 5 alternatives are available initially, thus $B = \{1,2,3,4,5\}$, with $v_5\left(w_5^0, y^0\right) = -\infty$. This implies that $h_5\left(w_5^0, y^0, w_5, y\right) = v_5\left(w_5, y\right)$, for all y. Therefore, a straight forward extension of (22) and (23) with $I_5\left(w_5^0, y^0, w_5, y\right) = 0$, yields

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right)$$

(25)
$$= \frac{\sum_{i=1}^{2} I_i\left(w_i^0, y^0, w_i, y\right) e^{v_i\left(w_i^0, y^0\right)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta} + e^{v_5(w_5, y)/\theta}\right)^{\theta-1} \sum_{i=3}^{4} I_i\left(w_i^0, y^0, w_i, y\right) e^{v_i\left(w_i^0, y^0\right)/\theta}}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta} + e^{v_5(w_5, y)/\theta}\right)^{\theta}}$$

for $y > 0$, and

(26)
$$E\, Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) = \sum_{i=1}^{2} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \frac{e^{v_i\left(w_i^0, y^0\right)}}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta} + e^{v_5(w_5, y)/\theta}\right)^{\theta}} + \sum_{i=3}^{4} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \frac{e^{v_i\left(w_i^0, y^0\right)/\theta}\left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta} + e^{v_5(w_5, y)/\theta}\right)^{\theta-1} dy}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{h_4(y)/\theta} + e^{v_5(w_5, y)/\theta}\right)^{\theta}}.$$

Example 3

In this example the initial setting is as in Example 1, while in the current period alternative 4 has been removed and is no longer available. Similarly to Example 2 this case is dealt with by letting $B = \{1, 2, 3, 4\}$, $v_4\left(w_4, y\right) = -\infty$ so that $h_4(y) = v_4\left(w_4^0, y^0\right)$, and $I_4\left(w_4^0, y^0, w_4, y\right) = 1$ for all y. As a result we obtain from (22) and (23) that

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right)$$

(27)
$$= \frac{\sum_{i=1}^{2} I_i\left(w_i^0, y^0, w_i, y\right) e^{v_i\left(w_i^0, y^0\right)} + \left(e^{h_3(y)/\theta} + e^{v_4\left(w_4^0, y^0\right)/\theta}\right)^{\theta-1}\left(I_3\left(w_3^0, y^0, w_3, y\right) e^{v_3\left(w_3^0, y^0\right)/\theta} + e^{v_4\left(w_4^0, y^0\right)/\theta}\right)}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{v_4\left(w_4^0, y^0\right)/\theta}\right)^{\theta}}$$

for $y > 0$, and

(28)
$$E\, Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) = \sum_{i=1}^{2} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \frac{e^{v_i\left(w_i^0, y^0\right)} dy}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{v_4\left(w_4^0, y^0\right)/\theta}\right)^{\theta}} + \sum_{i=3}^{4} \int_0^{y_i\left(w_i^0, y^0, w_i\right)} \frac{\left(e^{h_3(y)/\theta} + e^{v_4\left(w_4^0, y^0\right)/\theta}\right)^{\theta-1} e^{v_i\left(w_i^0, y^0\right)/\theta} dy}{e^{h_1(y)} + e^{h_2(y)} + \left(e^{h_3(y)/\theta} + e^{v_4\left(w_4^0, y^0\right)/\theta}\right)^{\theta}}$$

where $y_4\left(w_4^0, y^0, w_4\right)=\infty$.

Example 4

Suppose that the error terms of the utility function are i.i.d. with extreme value c.d.f., and

$$v_j\left(w_j, y\right)=\gamma\left(y-w_j\right)+a_j$$

where $\gamma>0$ is a parameter and a_j is a term that is supposed to capture non-pecuniary aspects of alternative j. In this case it is well known that

$$EY_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)=y_0+\frac{1}{\gamma}\log\left(\sum_{k\in B} e^{-\gamma w_k^0+a_k^0}\right)-\frac{1}{\gamma}\log\left(\sum_{k\in B} e^{-\gamma w_k+a_k}\right). \tag{29}$$

We shall now use the result of Corollary 4 to compute the c.d.f. of $Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)$.

$$\begin{aligned} & P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)>y\right) \\ & =\frac{\sum_{i\in B} I_i\left(w_i^0, y^0, w_i, y\right) e^{\gamma\left(y^0-w_i^0\right)+a_i^0}}{\sum_{i\in B} I_i\left(w_i^0, y^0, w_i, y\right) e^{\gamma\left(y^0-w_i^0\right)+a_i^0}+e^{\gamma y}\sum_{i\in B}\left(1-I_i\left(w_i^0, y^0, w_i, y\right)\right) e^{-\gamma w_i+a_i}} \end{aligned} \tag{30}$$

where

$$I_i\left(w_i^0, y^0, w_i, y\right)=\begin{cases}1 & \text{if} \quad y<y_i\equiv y^0+w_i-w_i^0+\dfrac{a_i^0-a_i}{\gamma} \\ 0 & \text{otherwise}.\end{cases}$$

Consider finally (30) in the special case when $B=\{1,2,3\}$. Suppose that $y_2<y_1<y_3$. Then it follows that $I_i\left(w_i^0, y^0, w_i, y\right)=0$, for $i=1,2,3,$ when $y\geq y_3$ and $I_i\left(w_i^0, y^0, w_i, y\right)=1$, for $i=1,2,3,$ when $y<y_2$. Therefore

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)>y\right)=0, \text{ for } y\geq y_3, \tag{31}$$

$$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right)>y\right)=1, \text{ for } y<y_2,$$

(32) $$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right) = \frac{e^{\gamma y^0 - \gamma w_3^0 + a_3^0}}{e^{\gamma y^0 - \gamma w_3^0 + a_3^0} + e^{\gamma y}\left[e^{a_1 - \gamma w_1} + e^{a_2 - \gamma w_2}\right]}$$

when $y_1 < y < y_3$, and

(33) $$P\left(Y_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) > y\right) = \frac{e^{\gamma y^0 - \gamma w_1^0 + a_1^0} + e^{\gamma y^0 - \gamma w_3^0 + a_3^0}}{e^{\gamma y^0 - \gamma w_1^0 + a_1^0} + e^{\gamma y^0 - \gamma w_3^0 + a_3^0} + e^{\gamma y - \gamma w_2 + a_2}}$$

when $y_2 \le y < y_1$.

Let us finally check that we get (27) when we use (20) and (31) to (33). For notational simplicity let $b_j = a_j - \gamma w_j$ and $b_j^0 = a_j^0 - \gamma w_j^0$. Then we have that $\gamma y_j = \gamma y^0 + b_j^0 - b_j$, and

$$EY_B\left(\mathbf{w}, V_B\left(\mathbf{w}^0, y^0\right)\right) = y_2 + \int_{y_2}^{y_1} \frac{\left(e^{\gamma y^0 + b_1^0} + e^{\gamma y^0 + b_3^0}\right)dy}{e^{\gamma y^0 + b_1^0} + e^{\gamma y^0 + b_3^0} + e^{\gamma y + b_2}} + \int_{y_1}^{y_3} \frac{e^{\gamma y^0 + b_3^0}dy}{e^{\gamma y^0 + b_3^0} + e^{\gamma y}\left(e^{b_1} + e^{b_2}\right)}.$$

By multiplying the numerator and the denominator of the respective integrands by $e^{-\gamma y}$ we realize that integration is immediate and the last expression reduces to

$$\begin{aligned}
& y_2 - \frac{1}{\gamma}\Big|_{y_2}^{y_1} \log\left(\left(e^{\gamma y^0 + b_1^0} + e^{\gamma y^0 + b_3^0}\right)e^{-\gamma y} + e^{b_2}\right) - \frac{1}{\gamma}\Big|_{y_1}^{y_3} \log\left(e^{\gamma y^0 + b_3^0 - \gamma y} + e^{b_1} + e^{b_2}\right) \\
&= y_2 - \frac{1}{\gamma}\log\left(e^{b_1^0 + \gamma\left(y^0 - y_1\right)} + e^{b_3^0 + \gamma\left(y^0 - y_1\right)} + e^{b_2}\right) + \frac{1}{\gamma}\log\left(e^{b_1^0 + \gamma\left(y^0 - y_2\right)} + e^{b_3^0 + \gamma\left(y^0 - y_2\right)} + e^{b_2}\right) \\
&- \frac{1}{\gamma}\log\left(e^{b_3^0 + \gamma\left(y^0 - y_3\right)} + e^{b_1} + e^{b_2}\right) + \frac{1}{\gamma}\log\left(e^{b_3^0 + \gamma\left(y^0 - y_1\right)} + e^{b_1} + e^{b_2}\right) \\
&= y^0 + \frac{b_2^0 - b_2}{\gamma} - \frac{1}{\gamma}\log\left(e^{b_3^0 + b_1 - b_1^0} + e^{b_1} + e^{b_2}\right) + \frac{1}{\gamma}\log\left(e^{b_3^0 + b_2 - b_2^0} + e^{b_3^0 + b_2 - b_2^0} + e^{b_2}\right) \\
&- \frac{1}{\gamma}\log\left(e^{b_3} + e^{b_1} + e^{b_2}\right) + \frac{1}{\gamma}\log\left(e^{b_3^0 + b_1 - b_1^0} + e^{b_1} + e^{b_2}\right) \\
&= y^0 + \frac{b_2^0 - b_2}{\gamma} + \frac{1}{\gamma}\log\left(\left(e^{b_1^0} + e^{b_2^0} + e^{b_3^0}\right)e^{b_2 - b_2^0}\right) - \frac{1}{\gamma}\log\left(e^{b_1} + e^{b_2} + e^{b_3}\right) \\
&= y^0 + \frac{1}{\gamma}\log\left(e^{b_1^0} + e^{b_2^0} + e^{b_3^0}\right) - \frac{1}{\gamma}\log\left(e^{b_1} + e^{b_2} + e^{b_3}\right)
\end{aligned}$$

which is equal to (29).

9. Conclusion

In this paper we have demonstrated that the notion of random expenditure function and compensated choice probabilities can be readily adapted within a discrete choice setting. We have moreover derived convenient analytic formulae for the Hicksian choice probabilities and the distribution of the Expenditure function and Compensating Variation. In particular we show that an aggregate version of Shephard's Lemma holds under rather general conditions. We have considered several special cases in detail. When the model belongs to the GEV class the formulae simplifies. As argued by McFadden (2001), this method it to be preferred to using simulations. Finally, we have discussed several examples. Although we have focused on Compensating Variation in this paper, the derivation of the distribution of Equivalent Variation is completely analogous.

Footnotes:

[1] Karlström (1998) and Dagsvik (2001) have independently obtained many of the same results as in this paper, although the proofs are different.

[2] See McFadden (1978).

[3] Note that the present notation accommodates the specification $v_j(w_j, y) = \psi_j(y - w_{1j}, w_{2j})$ for some function ψ_j that may depend on j, w_{2j} represents non-pecuniary attributes and w_{1j} the price (or user cost) associated with alternative j.

[4] In addition G^S must satisfy a number of regularity conditions to ensure that $\exp(-G^S(x_1,...,x_M))$ is a proper distribution function, cf. McFadden (1978).

[5] Here we adopt the definition of type I, II and III extreme value distributions used by Resnick (1987). Other authors use different enumerations.

References

AABERGE, R., J.K. DAGSVIK, and S. STRØM (1995), "Labor Supply Responses and Welfare Effects of Tax Reforms", *Scandinavian Journal of Economics*, **97**, 635-659.

CARNEIRO, P., K. HANSEN and J. HECKMAN (2001), "Removing the Veil of Ignorance in Assessing the Distributional Impacts of Social Policies", *Swedish Economic Policy Review*, **8**, 303-305.

DAGSVIK, J.K. (1994), "Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes", *Econometrica*, **62**, 1179-1205.

DAGSVIK, J.K. (1995), "How Large is the Class of Generalized Extreme Value Models?", *Journal of Mathematical Psychology*, **39**, 90-98.

DAGSVIK, J.K. (2001), "Compensated Variation in Random Utility Models" (Discussion Papers, no. 299, Statistics Norway).

DAGSVIK, J.K. (2002), "Discrete Choice in Continuous Time: Implications of an Intertemporal Version of the IIA Property", *Econometrica*, **70**, 817-831.

DIAMOND, P.A. and D. McFADDEN (1994), "Some uses of the Expenditure Function in Public Finance", *Journal of Public Economics*, **3**, 3-21.

DIEWERT, E. (1974), "Application of Duality Theory", in M.D. Intrilligator and D.A. Kendrick (eds.), *Frontier of Quantitative Economics*, Vol. II (North-Holland, Amsterdam).

HAMMOND, P. J. (1990), "Theoretical Progress in Public Economics: A Provocative Assessment", *Oxford Economic Papers*, **42**, 6-33.

HECKMAN, J. and B. HONORÉ (1990), "The Empirical Content of the Roy Model", *Econometrica*, **58**, 1121-1149.

HERRIGES, J.A. and C.L. KLING (1999), "Nonlinear Income Effects in Random Utility Models", *The Review of Economics and Statistics*, **81**, 62-72.

JOE, H. (2001), "Multivariate Extreme Value Distributions and Coverage of Ranking Probabilities", *Journal of Mathematical Psychology*, **45**, 180-188.

KARLSTRÖM, A. (1998), "Hicksian Welfare Measures in a Nonlinear Random Utility Framework" (Working paper, Royal Institute of Technology, Department of Infrastructure and Planning, Stockholm, presented at the 1st KFB Transport Economics Conference in Borlange, Sweden, June 1998).

KARLSTRÖM, A. (2001), "Welfare Evaluations in Non-Linear Random Utility Models with Income Effects", in D. Hensher (ed.), *Travel Behaviour Research: The Leading Edge* (Pergamon Press, Sydney).

LINDBERG, P.O., E. A. ERIKSSON and L.-G. MATTSSON (1995), "Invariance of Achieved Utility in Random Utility Models", *Environment and Planning* A, **27**, 121-142.

McFADDEN, D. (1978a), “Modelling the Choice of Residental Location”, in A. Karlqvist, L. Lundquist, F. Snickars and J.J. Weibull (eds.), *Spatial Interaction Theory and Planning Models* (North Holland, Amsterdam).

McFADDEN, D. (1978b), “Cost, Revenue, and Profit Functions”, in M. Fuss and D. McFadden (eds.), *Production Economics: A Dual Approach to Theory and Applications. Vol. I: The Theory of Production* (North-Holland, Amsterdam).

McFADDEN, D. (1981), “Econometric Models of Probabilistic Choice”, in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data* (MIT Press, London).

McFADDEN, D. (1999), “Computing Willingness-to-Pay in Random Utility Models”, in J. Moore, R. Riezman and J. Melvin (eds.), *Trade Theory and Econometrics: Essays in Honour of John S. Chipman* (Routledge, London).

McFADDEN, D. (1999), “Rationality for Economists?” *Journal of Risk and Uncertainty*, **19**, 73-105.

McFADDEN, D. and K. TRAIN (2000), “Mixed MNL Models for Discrete Response”, *Journal of Applied Econometrics*, **15**, 447-470.

McFADDEN, D. (2001), “Disaggregate Behavioural Travel Demand’s RUM Side: A 20-Year Retrospective”, in D. Hensher (ed.), *Travel Behaviour Research: The Leading Edge* (Pergamon Press, Sydney).

McKENZIE, L. (1957), “Demand Theory without a Utility Index” *Review of Economic Studies*, **24**, 185-189.

MOREY, E., R.D. ROWE and M. WATSON (1993), “A Repeated Nested-logit Model of Atlantic Salmon Fishing”, *American Journal of Agricultural Economics*, **75**, 578-592.

RESNICK, S.I. (1987), *Extreme Values, Regular Variation and Point Processes* (Springer-Verlag, New York).

ROY, R. (1947), “La Distribution du Revenue entre les divers Biens”, *Econometrica*, **15**, 205-225.

RUDIN, W. (1976), *Principles of Mathematical Analyis* (McGraw-Hill Inc., London).

SMALL, K.A. and H.S. ROSEN (1981), “Applied Welfare Economics with Discrete Choice Models”, *Econometrica*, **49**, 105-130.

STRAUSS, D.J. (1979), “Some Results on Random Utility Models”, *Journal of Mathematical Psychology*, **20**, 35-52.

VARIAN, H. (1992), *Microeconomic Analysis* (W.W. Norton & Co. Inc., New York).

Appendix

Proof of Theorem 3:

Recall first that

$$\{Y_B(\mathbf{w},u) > y\} \Leftrightarrow \{V_B(\mathbf{w},y) < u\}$$

which implies that

$$\{Y_B(\mathbf{w}, V_B(\mathbf{w}^0, y^0)) > y\} \Leftrightarrow \{V_B(\mathbf{w},y) < V_B(\mathbf{w}^0, y^0)\}.$$

If $J_B(w^0, y^0) = i$, the last event implies that $v_i(w_i^0, y^0) > v_i(w_i, y)$. Moreover,

$$\begin{aligned}
&P\left(Y_B(w, V_B(w^0, y^0)) > y, J_B(w^0, y^0) = i\right) \\
&= P\left(V_B(w,y) < V_B(w^0, y^0) = v_i(w_i^0, y^0) + \varepsilon_i\right) \\
&= P\left(v_i(w_i^0, y^0) + \varepsilon_i \geq \max\left(\max_{k\in B\backslash\{i\}}\left(v_k(w_k^0, y^0) + \varepsilon_k\right), \max_{k\in B\backslash\{i\}}\left(v_k(w_k, y) + \varepsilon_k\right)\right)\right) \\
&= P\left(v_i(w_i^0, y^0) + \varepsilon_i \geq \max_{k\in B\backslash\{i\}}\left(h_k(w_k^0, y^0, w_k, y) + \varepsilon_k\right)\right) \\
&= P\left(h_i(w_i^0, y^0, w_i, y) = \max_{k\in B}\left(h_k(w_k^0, y^0, w_k, y) + \varepsilon_k\right)\right) \\
&= \int F_i^B\left(u - h_1(w_1^0, y^0, w_1, y), u - h_2(w_2^0, y^0, w_2, y), \ldots, u - h_m(w_m^0, y^0, w_m, y)\right) du.
\end{aligned}$$

This completes the proof.

Q.E.D.

Proof of Theorem 4:

Consider the event

$$\begin{aligned}
&\left\{Y_B(\mathbf{w}, V_B(\mathbf{w}^0, y^0)) = y, J_B(\mathbf{w}^0, y^0) = i, J_B^*(\mathbf{w}^0, y^0, \mathbf{w}) = j\right\} \\
\Leftrightarrow &\left\{v_j(w_j, y) + \varepsilon_j = V_B(w, y) = V_B(w^0, y^0) = v_i(w_i^0, y^0) + \varepsilon_i\right\} \\
\Leftrightarrow &\left\{v_j(w_j, y) + \varepsilon_j = v_i(w_i^0, y^0) + \varepsilon_i, \max_{k\in B\backslash\{i\}}\left(v_k(w_k^0, y^0) + \varepsilon_k\right) \leq v_i(w_i^0, y^0) + \varepsilon_i, \max_{k\in B\backslash\{j\}}\left(v_k(w_k, y) + \varepsilon_k\right) \leq v_j(w_j, y) + \varepsilon_j\right\}.
\end{aligned}$$

This implies that $v_j(w_j, y) + \varepsilon_j \geq v_j(w_j^0, y^0) + \varepsilon_j$ and $v_i(w_j^0, y^0) + \varepsilon_i > v_i(w_i, y) + \varepsilon_i$ which is equivalent to $v_j(w_j, y) \geq v_j(w_j^0, y^0)$ and $v_i(w_i, y) < v_i(w_i^0, y^0)$.

It follows next that for small Δy

(A.1)

$$P\Big(Y_B\big(\mathbf{w}, V_B(\mathbf{w}^0, y^0)\big) \in (y, y+\Delta y), J_B(\mathbf{w}^0, y^0) = i, J_B^*(\mathbf{w}^0, y^0, \mathbf{w}) = j\Big)$$

$$= P\Big(v_i(\mathbf{w}^0, y^0) + \varepsilon_i = V_B(\mathbf{w}^0, y^0), v_j(w_j, y) + \varepsilon_j \le V_B(\mathbf{w}^0, y^0) \le v_j(w_j, y+\Delta y) + \varepsilon_j, \max_{k \in B\setminus\{j\}}(v_k(w_k, y) + \varepsilon_k) \le V_B(\mathbf{w}^0, y^0)\Big)$$

$$+ o(\Delta y)$$

$$= P\Big(v_i(w_i^0, y^0) + \varepsilon_i = V_B(\mathbf{w}^0, y^0), v_j(w_j, y) + \varepsilon_j \le V_B(\mathbf{w}^0, y^0), \max_{k \in B\setminus\{j\}}(v_k(w_k, y) + \varepsilon_k) \le V_B(\mathbf{w}^0, y^0)\Big)$$

$$-P\Big(v_i(w_i^0, y^0) + \varepsilon_i = V_B(\mathbf{w}^0, y^0), v_j(w_j, y+\Delta y) + \varepsilon_j \le V_B(\mathbf{w}^0, y^0), \max_{k \in B\setminus\{j\}}(v_k(w_k, y) + \varepsilon_k) \le V_B(\mathbf{w}^0, y^0)\Big) + o(\Delta y)$$

$$= P\Big(v_i(w_i^0, y^0) + \varepsilon_i \ge \max_{k \in B\setminus\{i\}}(v_k(w_k^0, y^0) + \varepsilon_k), v_j(w_j, y) + \varepsilon_j \le v_i(w_i^0, y^0) + \varepsilon_i, v_i(w_i^0, y^0) + \varepsilon_i \ge \max_{k \in B\setminus\{j\}}(v_k(w_k, y) + \varepsilon_k)\Big)$$

$$-P\Big(v_i(w_i^0, y^0) + \varepsilon_i \ge \max_{k \in B\setminus\{i\}}(v_k(w_k^0, y^0) + \varepsilon_k), v_j(w_j, y+\Delta y) + \varepsilon_j \le v_i(w_i^0, y^0) + \varepsilon_i, v_i(w_i^0, y^0) + \varepsilon_i \ge \max_{k \in B\setminus\{j\}}(v_k(w_k, y) + \varepsilon_k)\Big)$$

$$+ o(\Delta y).$$

Now recall that since $i \in B \setminus C(\mathbf{w}^0, y^0, \mathbf{w}, y)$ and $j \in C(\mathbf{w}^0, y^0, \mathbf{w}, y)$ we have that $v_j(w_j, y) = h_j(w_j^0, y^0, w_j, y)$ and $v_i(w_i^0, y^0) = h_i(w_i^0, y^0, w_i, y)$. As a result the probabilities above can be written as

(A.2)

$$P\Big(h_i(w_i^0, y^0, w_i, y) + \varepsilon_i = \max_{k \in B}\big(h_k(w_k^0, y^0, w_k, y) + \varepsilon_k\big)\Big)$$

$$-P\Big(h_i(w_i^0, y^0, w_i, y) + \varepsilon_i \ge v_j(w_j, y+\Delta y) + \varepsilon_j, h_i(w_i^0, y^0, w_i, y) + \varepsilon_i \ge \max_{k \in B\setminus\{j\}}\big(h_k(w_k^0, y^0, w_k, y) + \varepsilon_k\big)\Big)$$

$$+ o(\Delta y).$$

We realize that the last expression has the structure as the difference of two choice probabilities (apart from $o(\Delta y)$). Hence, (A.3) can be written as

$$\int F_i^B\big(u - h_1(w_1^0, y^0, w_1, y), u - h_2(w_2^0, y^0, w_2, y), \ldots, u - h_m(w_m^0, y^0, w_m, y)\big)du$$

$$-\int F_i^B\big(u - h_1(w_1^0, y^0, w_1, y), \ldots, u - h_{j-1}(w_{j-1}^0, y^0, w_{j-1}, y), u - v_j(w_j, y+\Delta y), \ldots, u - h_m(w_m^0, y^0, w_m, y)\big)du + o(\Delta y)$$

$$= \big(v_j(w_j, y+\Delta y) - v_j(w_j, y)\big)\int F_{ij}^B\big(u - h_1(w_1^0, y^0, w_1, y), u - h_2(w_2^0, y^0, w_2, y), \ldots, u - h_m(w_m^0, y^0, w_m, y)\big)du$$

which yields

(A.3)
$$P\left(Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)\in dy, J_B\left(\mathbf{w}^0,y^0\right)=i, J_B^*\left(\mathbf{w}^0,y^0,\mathbf{w}\right)=j\right)$$
$$=v_j\left(w_j,dy\right)\int F_{ij}^B\left(u-h_1\left(w_1^0,y^0,w_1,y\right),u-h_2\left(w_2^0,y^0,w_2,y\right),...,u-h_m\left(w_m^0,y^0,w_m,y\right)\right)du$$

for $i\in B\setminus C\left(\mathbf{w}^0,y^0,\mathbf{w},y\right)$ and $j\in C\left(\mathbf{w}^0,y^0,\mathbf{w},y\right)$, and zero for $i\notin B\setminus C\left(\mathbf{w}^0,y^0,\mathbf{w},y\right)$, $j\notin C\left(\mathbf{w}^0,y^0,\mathbf{w},y\right)$, $i\neq j$.

For $j=i$ it follows that

(A.4)
$$P\left(Y_B\left(\mathbf{w},V_B\left(\mathbf{w}^0,y^0\right)\right)=y, J_B\left(\mathbf{w}^0,y^0\right)=J_B^*\left(\mathbf{w}^0,y^0,\mathbf{w}\right)=i\right)$$
$$=P\left(v_i\left(w_i^0,y^0\right)=v_i\left(w_i,y_{ii}\left(w_i^0,y^0,w_i\right)\right), v_i\left(w_i^0,y^0\right)+\varepsilon_i>\max_{k\in B\setminus\{i\}}\left(v_k\left(w_k^0,y^0\right)+\varepsilon_k\right)\right)$$
$$=P\left(J_B\left(\mathbf{w}^0,y^0\right)=i\right)$$

for $y=y_i\left(w_i^0,y^0,w_i\right)$. Moreover, the expenditure function in this case has measure zero outside the point $y_i\left(w_i^0,y^0,w_i\right)$.

Note finally that $i\in B\setminus C\left(\mathbf{w}^0,y^0,\mathbf{w},y\right)$ and $j\in C\left(\mathbf{w}^0,y^0,\mathbf{w},y\right)$ means that $v\left(w_i,y\right)<v\left(w_i^0,y^0\right)$ and $v\left(w_j,y\right)\geq v\left(w_j^0,y^0\right)$, which is equivalent to $y_j\left(w_j^0,y^0,w_j\right)\leq y<y_i\left(w_i^0,y^0,w_i\right)$.

This completes the proof.

Q.E.D.

Proof of Lemma 1:

Suppose first that

$$\int_0^\infty x^\alpha dH(x)<\infty .$$

Then for any $b\geq 0$

(A.5)
$$\int_b^\infty x^\alpha dH(x)\geq b^\alpha\int_b^\infty dH(x)=b^\alpha\left(1-H(b)\right).$$

When $b\to\infty$ the left hand side of (A.5) tends towards zero, which yields

(A.6) $$\lim_{b\to\infty} b^{\alpha}\left(1-H(b)\right)=0\,.$$

Using integration by parts we obtain that

(A.7) $$\int_0^b x^{\alpha} d\,H(x) + b^{\alpha}\left(1-H(b)\right) = \alpha\int_0^b x^{\alpha-1}\left(1-H(x)\right)dx\,.$$

When $b\to\infty$, (A.6) and (A.7) imply that

(A.8) $$\int_0^{\infty} x^{\alpha} d\,H(x) = \alpha\int_0^{\infty} x^{\alpha-1}\left(1-H(x)\right)dx$$

so that the right hand side of (A.9) also exists. Suppose next that

$$\int_0^{\infty} x^{\alpha} d\,H(x) = \infty\,.$$

Then (A.7) implies that also

$$\int_0^{\infty} x^{\alpha-1}\left(1-H(x)\right)dx = \infty\,.$$

Hence we have proved that (A.8) holds in either case.

Q.E.D.

Proof of Corollary 3:

From Theorem 4 and homogeneous property of $\log F^{B}$ when F^{B} is a multivariate extreme value distribution we get, with $h_j = h_j\left(w_j^0, y^0, w_j, y\right)$,

$$\begin{aligned}
&\int F_{ij}^{B}\left(u-h_1, u-h_2,..., u-h_m\right)du \\
&= -\int \exp\left(-e^{-u}G^{B}\left(-h_1,-h_2,...,-h_m\right)\right)\left(e^{-2u}G_i^{B}\left(-h_1,-h_2,...,-h_m\right)G_j^{B}\left(-h_1,-h_2,...,-h_m\right) - e^{-u}G_{ij}^{B}\left(-h_1,-h_2,...,-h_m\right)\right)du \\
&= -\frac{G_i^{B}\left(-h_1,-h_2,...,-h_m\right)G_j^{B}\left(-h_1,-h_2,...,-h_m\right)}{G^{B}\left(-h_1,-h_2,...,-h_m\right)^2} + \frac{G_{ij}^{B}\left(-h_1,-h_2,...,-h_m\right)}{G^{B}\left(-h_1,-h_2,...,-h_m\right)}.
\end{aligned}$$

From this result the result of Corollary 3 follows.

Q.E.D.