Statistisk sentralbyrå
Statistics Norway

Marginal compensated effects and the slutsky equation for discrete choice models

TALL SOM FORTELLER

John K. Dagsvik

DISCUSSION PAPERS

930

Discussion Papers No. 930, May 2020
Statistics Norway, Research Department

John K. Dagsvik

Marginal compensated effects and the slutsky equation for discrete choice models

Abstract:
In many instances the consumer faces choice settings where the alternatives are discrete. Examples include choice between variants of differentiated products, urban transportation modes, residential locations, types of education, etc. So far, a Slutsky equation for discrete choice models has not been derived. In this paper an aggregate Slutsky equation for the discrete case is obtained, which differs in important ways from the corresponding equation in the standard theory of consumer demand. A remarkable feature of the compensated marginal effects in the discrete case is that they are usually not symmetric, as the marginal compensated effects with respect to a price increase versus a price decrease may be different. The description of the analytic formulas is accompanied by several examples of their use: for example, in travel demand and labor supply.

Keywords: Equivalent variation, Compensating variation, Discrete/continuous choice, Slutsky equation, Marginal compensated effects, Price indexes

JEL classification: C25, C43, D11

Acknowledgments: I am grateful for comments by Brita Bye, Thor O. Thoresen, Terje Skjerpen and Steinar Strøm, and by participants at seminars at the Research Department of Statistics Norway and the Frisch Centre of Economic Research. Part of this research has had financial support from the Research Council of Norway (the tax research program) and the Frisch Centre of Economic Research.

Address: PO Box 2633, St. Hanshaugen, NO-0131, Statistics Norway, Research Department. E-mail: john.dagsvik@ssb.no

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

I standard lærebokteori for konsumentenes valg antas godene å være tilgjengelige i uendelig delbare kvanta. I mange valgsituasjoner er imidlertid godene som etterspørres kvalitative (diskrete), slike som varianter av differensierte produkter (transportalternativer, biler, viner, type jobb, type utdanning, vaskemaskiner, etc.). For slike situasjoner kan diskret valghandlingsteori (diskrete valgmodeller) benyttes til å analysere valgatferden til konsumentene. I standard konsumentteori kan som kjent den såkalte Slutsky likningen benyttes til å beregne marginale kompenserte priseffekter. Slutsky likningen gir sammenhengen mellom de kompenserte og de ukompenserte pris- og inntektselastisitetene, slik at dersom en kjenner de marginal ukompenserte pris- og inntektseffektene kan en tallfeste de tilsvarende marginale kompenserte effektene.

I denne artikkelen utledes en aggregert Slutsky likning for diskrete valgmodeller. Det har tidligere ikke eksistert en tilsvarende Slutsky likning for slike modeller. Den diskrete Slutsky likningen gjør det dermed mulig å beregne kompenserte priselastisiteter for andelen (valgsannsynligheten) som etterspør et diskret alternativ på grunnlag av de tilsvarende ukompenserte pris og inntektselastisiteter. Slutsky-ligningen i dette tilfellet skiller seg på sentrale måter fra den tilsvarende ligningen i standard konsumentteori. For eksempel er venstre - og høyrederiverte av de kompenserte valgsannsynlighetene med hensyn på pris som regel forskjellige.

Til slutt diskuterer vi utvalgte spesialtilfeller. Ved hjelp av Slutsky likningen vises hvordan substitusjonseffekter i disse tilfellene kan beregnes via de kompenserte elastisitetene med hensyn på endringer i reisekostnader og reisetider.

1. Introduction

The theory of compensated demand and compensating and equivalent variations is well developed for the case when the commodity space is infinitely divisible: see, for example, Hausman (1981). However, in many instances the consumer faces choice settings where the alternatives are discrete. These include choice between variants of a differentiated product, urban transportation modes, residential locations, types of education, types of child care, etc. In the context of discrete choice settings, one cannot apply the standard microeconomic textbook approach to express demand functions. The reason is that the set of feasible consumption alternatives is not a continuum and the utility function is not differentiable. Thus, the standard approach based on marginal calculus does not apply.

As regards welfare analysis, the standard tools of applied welfare economics are not directly applicable in discrete choice situations. Nevertheless, it is important to develop practical welfare measures in these settings also, because welfare judgments are of major interest in several areas, such as the choice between transportation modes, housing alternatives, variants of differentiated products, types of schooling, and types of childcare. In these areas, welfare evaluations of public policies which change prices, taxes, and quality attributes of some alternatives are relevant.

A central aspect of welfare assessment is the calculation of marginal compensating effects. In the traditional approach to microeconomic analysis, with infinitely divisible quantities of goods the Slutsky equation plays a key role. The Slutsky equation, referred to as the "fundamental equation in value theory" by Hicks (1936), allows one to compute the compensating price elasticities from the corresponding uncompensated price and income elasticities. Specifically, marginal compensated (Hicksian) effects are used to justify key price indexes and they also play an important role in the analysis of optimal taxation. In the special case where the utility function is linear in income there are no income effects and the marginal compensated effects will in general be different from the corresponding uncompensated effects. In this case EV and CV can be readily expressed on closed form for (McFadden, 1999, and Niemeier, 1997). However, when utilities are non-linear in income one can no longer express these welfare measures and marginal compensated effects by simple formulas.

Dagsvik and Karlström (2005) obtained analytic formulas for compensated choice probabilities and the distribution of welfare measures such as CV and EV in discrete choice models when utility is non-linear in income.[1] In this paper, we employ the results obtained by Dagsvik and Karlström (2005) to derive compensated marginal effects for discrete choice models and to establish a corresponding discrete Slutsky equation (discrete Slutsky equation). This Slutsky equation also covers a specific case of discrete/continuous choice. It turns out that the discrete Slutsky equation is in part analogous to the standard Slutsky equation, but also differs in essential ways. A remarkable feature of the compensated marginal effects in the discrete case is that they are usually not symmetric, as the marginal compensated effects with respect to a price increase versus a price decrease may be different. In a separate paper (Dagsvik et al., 2019), marginal compensated effects for discrete labor supply models are analyzed.

An early general and seminal treatment of welfare analysis in discrete/continuous choice models was undertaken by Small and Rosen (1981). They seem to be the only ones who have previously discussed marginal compensated effects. Their treatment is, however, incomplete and seems partly misleading, as will be discussed further below.

The paper is organized as follows. In Section 2 we discuss the notion of compensating choice and the random expenditure function in the discrete choice setting. Section 3 deals with joint ex-ante and ex-post compensated choice. In Section 4 we discuss marginal compensated effects with special reference to the Slutsky equation and in Section 5 we discuss some selected examples.

2. Compensated discrete choice

In discrete choice theory based on random utility representations, the notion of compensated demand is more complicated than in the conventional case. Also, separate treatments are necessary for the one-period setting and the two-period setting: that is, before (ex-ante) and after (ex-post) a reform is introduced. The reason for this is that in random utility models there is no unique deterministic correspondence between prices, expenditure, and utility because (indirect) utility is a random function of prices and income. The random utility

---

[1] Kornstad and Thoresen (2006) and Dagsvik et al. (2009) have conducted welfare analyses based on the welfare measures derived in Dagsvik and Karlström (2005).

representation is motivated by the fact that not all variables that influence preferences are observable to the researcher. Some of the variables that affect preferences may even be random to the consumer himself. The reason is that tastes may vary in an unpredictable way from one moment to the next across identical choice settings due to psychological factors and difficulties with making a definitive assessment of the value of the alternatives. Consequently, the utility function, the expenditure function, CV, and EV all become interdependent random variables. This feature calls for a careful probabilistic analysis in the derivation of the respective distribution functions.

Consider a general setting where the consumer faces a choice between a composite continuous good and a set of discrete alternatives where the discrete alternatives are mutually exclusive. Let $U^*(x_0, x_j)$ be the utility of the quantity x_0 of the composite good and the quantity x_j associated with the discrete alternative *j*, *j* = 1, 2,…, *m*. Most of the time x_j will be equal to 1 (when alternative *j* is chosen) or zero, but for the sake of comparison with Small and Rosen (1981) we shall also consider briefly the discrete/continuous case in which x_j takes values in $(0,\infty)$. The consumer maximizes $U_j^*(x_0, x_j)$ subject to the budget constraint

$$x_0 + \sum_{j=1}^{m} x_j w_j = y, \;\; x_j \geq 0, \;\; x_j x_k = 0, \;\; k \neq j,$$

where *y* denotes income, w_j the price of the discrete alternative *j*, and the price of the composite indivisible good with quantity x_0 is normalized to 1. Let $U_j(w_j, y)$ be the conditional indirect utility given the discrete alternative *j*. That is, $U_j(w_j, y)$ is the maximum of $U_j^*(x_0, x_j)$ subject to the budget constraint $x_0 + x_j w_j = y.$ In the pure discrete case, where $x_j = 1$ the indirect utility conditional on alternative *j* admits the form $U_j(w_j, y) = U_j^*(y - w_j, 1).$ The general formulation above covers several cases (but not all) of interest. Consider, for example, the choice of working in different labor market sectors, where it is understood that hours of work are fixed and possibly sector-specific. In this case of sectoral labor supply without taxes and with fixed hours of work, the conditional indirect utility can be expressed as $U_j(w_j, y) = U_j^*(y + w_j, 1).$ Thus, the function $U_j(w_j, y)$ can be both increasing (occupational mobility and labor supply) and decreasing in w_j.

Assumption 1

The utility of alternative j has the structure $U_j(w_j, y) = v_j(w_j, y) + \varepsilon_j$, *where* $v_j(w_j, y)$ *is a deterministic function that is strictly increasing in y and strictly monotone in* w_j *and* $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_m)$ *is a stochastic vector with joint c.d.f. F that possesses a continuous probability density.*

As mentioned above, the stochastic terms are supposed to account for the effect on preferences from variables that are not observed by the researcher. Under Assumption 1 and if *F* is a multivariate extreme value distribution with Gumbel marginals, then the implied choice model becomes the Generalized Extreme Value (GEV) model (McFadden, 1978).

Recall that the additive random utility structure assumed above, which is the same set-up as in Dagsvik and Karlström (2005), represents no loss of generality. Bhattacharya (2015, 2018) shows that one can obtain formulas for the distribution of CV and EV without assuming a separable utility structure as in Assumption 1. However, as Dagsvik (1994, 1995) and Joe (2001) have demonstrated, any random utility model can be approximated arbitrarily closely by a GEV model. Since the GEV family is a subclass of the random utility models generated by Assumption 1 it follows that there is no loss of generality in postulating Assumption 1.

The agent's choice set of available alternatives may be a subset of the universal set of all possible alternatives. For simplicity, let {1, 2, …, *m*} denote the index of all possible alternatives. If alternative *j* is not available to the agent then the corresponding price, $w_j = \infty$ and $\partial v_j(w_j, y) / \partial w_j = 0$. Evidently, this represents no loss of generality. Let $J(w, y)$ be the (Marshallian) choice function and $w = (w_1, ..., w_m)$ the vector of prices. That is, $J(w, y) = j$ if the discrete alternative *j* is chosen, given prices and income (*w*, *y*). It follows from McFadden (1981, pp. 212–14) that under Assumption 1 the choice probabilities are given by

(2.1)
$$\varphi_j(w, y) = P(J(w, y) = j) = P(v_j(w_j, y) + \varepsilon_j = \max_{r \le m}(v_r(w_r, y) + \varepsilon_r))$$
$$= \int_{-\infty}^{\infty} F_j'(u - v_1(w_1, y), u - v_2(w_2, y)v_2, ..., u - v_m(w_m, y)))du.$$

We call $\varphi_j(w,y)$ the Marshallian (or uncompensated) choice probability of choosing alternative *j*. The corresponding conditional demand function given the discrete alternative *j* follows from Roy's identity, namely

(2.2) $$x_j = x_j(w_j, y) = -\frac{v'_{j1}(w_j, y)}{v'_{j2}(w_j, y)}$$

where v'_{jk}, *k* = 1, 2, denotes the derivative with respect to component *k*. We note that the conditional demand functions defined above are deterministic since they depend only on the deterministic terms of the utility function. More general utility specifications for discrete/continuous choice are given by Dubin and McFadden (1984), Hanemann (1984) and Dagsvik (1994).

Define next the conditional expenditure function $e_j(w_j, u)$, given alternative *j*, by the relation

$$u = U_j(w_j, e_j(w_j, u)) = v_j(w_j, e_j(w_j, u)) + \varepsilon_j$$

where *u* is a given utility level. When $v_j(w, y)$ is strictly increasing in *y* it follows that $e_j(w, u)$ is uniquely determined. The expenditure function (unconditional) $e(w,u)$ is therefore given by

$$e(w,u) = \min_{j \le m} e_j(w_j, u).$$

Since the utility function depends on random taste variables the expenditure function becomes stochastic. The Hicksian (or compensated) conditional demand x_j^H given alternative *j* equals

$$x_j^H = x_j^H(w_j, u) = x_j(w_j, e_j(w_j, u)).$$

Let $J^H(w,u)$ denote the Hicksian discrete choice function given prices and utility level (*w*, *u*). The concept that corresponds to (aggregate) Hicksian demand is the Hicksian (or compensated) choice probability. It is defined as

$$P_j^H(w,u) = P(J^H(w,u) = j) = P(e_j(w_j, u) = e(w,u)).$$

Dagsvik and Karlström (2005, Theorem 2) have derived the formula for $P_j^H(w,u)$ under Assumption 1. They also obtained a discrete version of Shephard's lemma for the standard discrete choice case and, furthermore, the distribution of the expenditure function (Theorem 1). Another way of expressing the Hicksian choice function is as $J^H(w,u) = J(w, e(w,u))$. If

$J(w,y)$ and $e(w,u)$ were independent random functions, then one could derive the Hicksian choice probability from the relation

$$P(J^H(w,u)=j)=P(J(w,e(w,u))=j)=\int_0^\infty P(J(w,y)=j)P(e(w,u)\in dy).$$

Unfortunately, the random functions $J(w,y)$ and $e(w,u)$ are in general stochastically interdependent and therefore the equation above is not always true. In fact, we have the following result.

Proposition 1

Assume that Assumption 1 holds. Then

$$P(J(w,y)=j\,|\,e(w,u)=y)=P(J(w,y)=j)$$

only if $\partial v_j(w_j,y)/\partial y$ *is independent of j and F is a multivariate extreme value c.d.f.*

The proof of Proposition 1 is given in the appendix. Proposition 1 shows that even if the choice model is a conditional logit model the deterministic part of the utility function must be linear in income with a coefficient associated with income which is independent of the alternatives.

3. Joint ex-ante choice and ex-post compensated choice

The focus of this paper is the analysis of compensated choice behavior in the two-period setting where the first period is called ex-ante (before the reform is introduced) and the second period ex-post (after the reform has been introduced). Let the income and price of alternative *j* ex-ante be equal to (y,w_j) and the price of alternative *j* ex-post be equal to $\tilde{w}_j$. Recall that we have adopted the convention that when an alternative *j* (say) does not belong to the choice set of the consumer, the corresponding price (relative to the individual) is equal to infinity and accordingly the corresponding utility is equal to minus infinity. It is assumed that the stochastic terms of the utility function are not affected by the reforms. This assumption is common in these types of welfare analysis and it simply means that the welfare effects are interpreted as conditional on all factors other than the actual prices (or wages) being kept fixed. Under Assumption 1 we shall, for simplicity, sometimes write v_j or $v_j(y)$ instead of $v_j(w_j, y)$. Let $V(w, y)$ be the (unconditional) ex-ante indirect utility functions defined by

(3.1) $$V(w, y) = \max_{j \le m} U_j(w_j, y).$$

Define Y_j by

$$Y_j = Y_j(\tilde{w}_j; w, y) = e_j(\tilde{w}_j, V(w, y))$$

for *j* belonging to the ex-post choice set and define Y by

$$Y = Y(\tilde{w}; w, y) = \min_{j \le m} Y_j(\tilde{w}_j; w, y).$$

Whereas Y_j is a conditional ex-post expenditure function given alternative *j*, Y is the unconditional ex-post expenditure function that yields the income required to maintain the ex-post utility level equal to the ex-ante utility level. The corresponding compensating variation measure is defined by $CV = Y - y$. Alternatively, Y can be obtained as the solution to the equation

$$V(w, y) = V(\tilde{w}_j, Y).$$

We have now defined the theoretical concepts that are necessary for deriving analytic results that are analogous to the one-period expenditure function and Hicksian demands. Let $Q^H(j, k, z; \tilde{w}, w, y)$ be the joint probability of choosing alternative *j* ex-ante, alternative *k* ex-post, and $\{Y \le z\}$ when the ex-post maximum utility is equal to the ex-ante maximum utility. Thus,

$$Q^H(j, k, z; \tilde{w}, w, z) = P(U_j(w_j, z) = \max_r U_r(w_r, y) = U_k(\tilde{w}_k, Y) = \max_r U_r(\tilde{w}_r, Y), Y \le z)$$

for $z \ge 0$. For notational simplicity we shall, most of the time, write $Q^H(j, k, z)$ instead of $Q^H(j, k, z; \tilde{w}, w, y)$. Let $Q^H(j, k) = Q^H(j, k, \infty)$, which is the joint compensated (Hicksian) probability of choosing alternative *j* ex-ante and alternative *k* ex-post (which means that the respective utility levels of the chosen alternatives before and after the reform are the same). Let

$$Q_j^H = Q_j^H(\tilde{w}, y, w) = P(Y_j(\tilde{w}_j; w, y) = Y(\tilde{w}; w, y))$$

which is the probability of choosing alternative *j* ex-post conditional on the ex-post utility level being equal to the ex-ante utility level. Evidently,

$$Q_j^H = \sum_{r>0} Q^H(r, j).$$

If $V(w,y)$ were exogenous one could obtain Q_j^H from $P_j^H(w,u)$ because in this case one would have that $Q_j^H(\tilde{w},w,y) = EP_j^H(\tilde{w},V(w,y)),$ where the expectation is taken with respect to $V(w,y).$ However, Proposition 1 implies that the latter equation *does not* hold because $V(w,y)$ and $\{\tilde{\varepsilon}_j(\tilde{w}_j,u)\}$ depend on the same random taste shifters $\{\varepsilon_j\}$ and in this sense $V(w,y)$ is endogenous.

4. The discrete Slutsky equation

We start with a brief review of the Slutsky relation in standard consumer theory. Let $d_j(p,y)$ denote the (Marshallian) demand of commodity *j* as a function of prices and income (p, y) and let $d_j^H(p,u)$ denote the corresponding Hicksian demand function where *u* is the utility level and $e(p,u)$ the corresponding expenditure function. The Hicksian demand function is not directly observable because it depends on the unobservable utility level. However, Slutsky (1915) showed how the marginal Hicksian demands can be obtained from the corresponding marginal Marshallian demands through the so-called Slutsky equation given by

(4.1) $$\partial d_j(p,e(p,u))/\partial p_k = \partial d_j^H(p,u)/\partial p_k - d_k(p,e(p,u))\partial d_j(p,e(p,u))/\partial y.$$

This equation allows one to compute the unobserved marginal compensated demands with respect to price changes by using the corresponding Marshallian demands (Varian, 1992).

Consider next the two-period discrete case. This is more complicated because the preferences are stochastic and the ex-ante indirect utility is endogenous, as discussed above. Define, for positive *j* and *k*,

$$\frac{\partial^+\varphi_k^H}{\partial w_j} = \lim_{\tilde{w}_k \downarrow 0} \frac{Q_k^H(\tilde{w},w,y) - \varphi_k(w,y)}{\tilde{w}_j - w_j} \quad \text{and} \quad \frac{\partial^-\varphi_k^H}{\partial w_j} = \lim_{\tilde{w}_k \uparrow 0} \frac{Q_k^H(\tilde{w},w,y) - \varphi_k(w,y)}{\tilde{w}_j - w_j}.$$

The expressions above are the right and left derivatives of $Q_k^H(\tilde{w},w,y)$ with respect to the ex-post price of alternative *j*. They correspond to the right and left marginal compensated effects of the choice probability of alternative *k* resulting from a price increase or a price decrease, respectively, of alternative *j*. As we shall see below, it turns out that in general one has $\partial^+\varphi_k/\partial w_j \neq \partial^-\varphi_k/\partial w_j$, which means that in general the derivative $\partial\varphi_j/\partial w_k$ does not exist.

Theorem 1 (discrete Slutsky equation)

Assume that Assumption 1 holds with $v_j(w_j, y)$ *that is continuously differentiable and decreasing in* w_j *for all j. Then*

$$\frac{\partial\varphi_j}{\partial w_j} = \frac{\partial^+\varphi_j^H}{\partial w_j} - x_j \cdot \frac{\partial\varphi_j}{\partial y}, \qquad \frac{\partial\varphi_j}{\partial w_j} = \frac{\partial^-\varphi_j^H}{\partial w_j},$$

$$\frac{\partial\varphi_k}{\partial w_j} = \frac{\partial^+\varphi_k^H}{\partial w_j} \cdot \frac{\partial v_j / \partial y}{\partial v_k / \partial y} \quad \text{and} \quad \frac{\partial\varphi_k}{\partial w_j} = \frac{\partial^-\varphi_k^H}{\partial w_j}$$

for $k \neq j,\ j,k > 0,$ *where*

$$x_j = -\frac{\partial v_j / \partial w_j}{\partial v_j / \partial y}.$$

The proof of Theorem 1 is given in the appendix. Recall that in the conventional continuous case the demand for good *j* is given by Roy's identity. Hence, x_j has the interpretation as the conditional demand, given the choice *j*. Accordingly, we realize that only the equation determining $\partial^+\varphi_j^H / \partial w_j$ is similar to the standard Slutsky equation in (4.1) with income effect given by $x_j \partial\varphi_j / \partial y$.

In many applications preferences are assumed to depend on non-pecuniary and alternative specific attributes. For example, in analysis of urban travel behavior attributes such as "on-vehicle times" and "out-of-vehicle times" play a major role. The results of Theorem 1 or Corollary 1 can also be applied to calculate marginal compensating effects with respect to selected non-pecuniary attributes simply by replacing w_j with the selected attribute in the formulas of Theorem 1/Corollary 1.

Surprisingly, the corresponding Slutsky equation for price decreases is different in that the marginal compensated price effect is equal to the marginal uncompensated price effect. In order to understand why, let us, for simplicity, consider the binary case with two alternatives. The argument in the general case with many alternatives is similar. Consider first the case where w_2 increases to $\tilde{w}_2 > w_2$, whereas w_1 remains unchanged. Then $Q^H(1,2) = 0$ because there is nothing to gain by switching from alternative 1 to alternative 2. Define y_2 by the relation $v_2(w_2, y) = v_2(\tilde{w}_2, y_2).$ Consider next the case where the agent chooses alternative 2 ex-ante and ex-post given that utility is kept constant. In this case expenditure Y is determined by $U_2(w_2, y) = U_2(\tilde{w}_2, Y)$, which implies that $Y = y_2$. We therefore obtain that

(4.2) $$Q^H(2,2) = P(U_1(w_1, y) < U_2(w_2, y) = U_2(\tilde{w}_2, y_2) > U_1(w_1, y_2)).$$

Evidently, $y_2 > y$, which implies that $U_1(w_1, y_2) > U_1(w_1, y)$. Consequently, (4.2) reduces to

$$Q^H(2,2) = P(U_2(\tilde{w}_2, y_2) > U_1(w_1, y_2))$$

so that

(4.3) $$Q_2^H(\tilde{w}, w, y) - \varphi_2(w, y) = Q^H(1,2) + Q^H(2,2) - \varphi_2(w, y) = Q^H(2,2) - \varphi_2(w, y)$$
$$= P(U_2(\tilde{w}_2, y_2) > U_1(w_1, y_2)) - \varphi_2(w, y) = \varphi_2(\tilde{w}, y_2) - \varphi_2(w, y)$$

where $\tilde{w} = (w_1, \tilde{w}_2)$. Consider next the case where $\tilde{w}_2 < w_2$. Then, evidently, $Q^H(2,2) = \varphi_2$ because there is nothing to gain by switching to alternative 1. Furthermore, if alternative 1 was chosen ex-ante the agent may switch to alternative 2 ex-post. This will happen if

$$\{U_2(w_2, y) < U_1(w_1, y) = U_2(\tilde{w}_2, Y) > U_1(w_1, Y)\}.$$

The latter event implies that

$$\{U_1(w_1, Y) < U_1(w_1, y), U_2(\tilde{w}_2, Y) > U_2(w_2, y)\} \Leftrightarrow \{v_1(w_1, Y) < v_1(w_1, y), v_2(\tilde{w}_2, Y) > v_2(w_2, y)\}$$

which is equivalent to $\{Y \in (y_2, y)\}$. Consequently, it follows that

(4.4) $$Q^H(1,2) = P(U_2(w_2, y) < U_1(w_1, y) = U_2(\tilde{w}_2, Y) > U_1(w_1, Y))$$
$$= P(U_1(w_1, y) = U_2(\tilde{w}_2, Y), Y \in (y_2, y)\} = P(U_2(w_2, y_2) < U_1(w_1, y) \le U_2(\tilde{w}_2, y)).$$

Accordingly, (4.4) yields

(4.5) $$Q_2^H(\tilde{w}, w, y) - \varphi_2(w, y) = Q^H(1,2) = P(U_2(w_2, y) \le U_1(w_1, y) \le U_2(\tilde{w}_2, y)) = \varphi_2(\tilde{w}, y) - \varphi_2(w, y).$$

We note that the final expressions in (4.3) and (4.5) differ in an important way. The expression in (4.5) is equal to the own marginal uncompensated price effect. In the former case in (4.3) the corresponding expression is similar apart from the fact that income is replaced by y_2. (For the sake of interpretation, note that by implicity differentiation and Roy's identity it follows that $\partial y_2 / \partial w_2 = x_2$). Thus, in the case of a price increase the own marginal compensated price effect is obtained by substituting income with y_2 in the formula for the corresponding marginal uncompensated price effect. This means that an income effect is present, represented by y_2. This is due to the fact that when the price of alternative 2 increases the marginal own compensated price effect equals $Q^H(2,2) - \varphi_2$, whereas the marginal own compensated price effect equals $Q^H(1,2)$ when the price of alternative 2 decreases.

The asymmetry in the Slutsky equation is not restricted to the case where utility is additively separable in a deterministic and a stochastic term. The essential difference from the separable case is that in the non-separable case y_2 may be stochastic and determined by $U_2(w_2, y, \varepsilon_2) = U_2(\tilde{w}_2, y_2, \varepsilon_2)$. Still the argument above, with minor modification, goes through. Section 5 we demonstrate that the asymmetry in the Slutsky equation also may occur in the standard textbook model of labor force participation in the presence of non-linear taxes.

We note above that in some cases, such as in models of labor suppy and occupational mobility, the utility function is increasing in prices (wage rates). By symmetry the result in next corollary follows readily from Theorem 1.

Corollary 1

Assume that Assumption 1 holds with $v_j(w_j, y)$ that is continuously differentiable and increasing in w_j for all j. Then

$$\frac{\partial \varphi_j}{\partial w_j} = \frac{\partial^- \varphi_j^H}{\partial w_j} + x_j \cdot \frac{\partial \varphi_j}{\partial y}, \qquad \frac{\partial \varphi_j}{\partial w_j} = \frac{\partial^+ \varphi_j^H}{\partial w_j},$$

$$\frac{\partial \varphi_k}{\partial w_j} = \frac{\partial^- \varphi_k^H}{\partial w_j} \cdot \frac{\partial v_j / \partial y}{\partial v_k / \partial y} \quad \textit{and} \quad \frac{\partial \varphi_k}{\partial w_j} = \frac{\partial^+ \varphi_k^H}{\partial w_j}$$

for $k \neq j,\ j,k > 0,$ *where*

$$x_j = \frac{\partial v_j / \partial w_j}{\partial v_j / \partial y}.$$

From Theorem 1 the next corollary is also immediate.

Corollary 2

Assume that Assumption 1 holds with $v_j(w_j, y) = \gamma(y - w_j)$ for some function $\gamma(y)$ that is independent of j. Then

$$\frac{\partial \varphi_j}{\partial w_j} = \frac{\partial^+ \varphi_j^H}{\partial w_j} - \frac{\partial \varphi_j}{\partial y}, \quad \frac{\partial \varphi_j}{\partial w_j} = \frac{\partial^- \varphi_j^H}{\partial w_j},$$

$$\frac{\partial \varphi_k}{\partial w_j} = \frac{\partial^+ \varphi_k^H}{\partial w_j} \quad \textit{and} \quad \frac{\partial \varphi_k}{\partial w_j} = \frac{\partial^- \varphi_k^H}{\partial w_j}$$

for $k \neq j,\ j,k > 0.$

As mentioned above, the case of discrete labor supply models (e.g. van Soest, 1995, and Dagsvik and Strøm, 2006) is analyzed in a separate paper (Dagsvik et al., 2019). The reason is that this case does not immediately fit into the framework considered here because the price (wage rate) is the same for all alternatives (different hours of work schedules).

Consider next the discrete/continuous case where the conditional demand functions are determined by Roy's identity, as explained above. Let $\bar{X}_j = x_j \varphi_j$, where x_1 is given by (2.2). That is, $\bar{X}_j$ is the unconditional (aggregate) demand for alternative *j* in the case of discrete/continuous choice. For the conditional demands the direct marginal effect must obviously satisfy the conventional Slutsky equation: that is,

$$\frac{\partial x_j}{\partial w_j} = \frac{\partial x_j^H}{\partial w_j} - x_j \cdot \frac{\partial x_j}{\partial y}.$$

Furthermore, since x_k does not depend on w_j for $k \neq j$, it follows that $\partial x_k^H / \partial w_j = 0.$ We therefore obtain the relation

$$\frac{\partial^{\pm} \bar{X}_j^H}{\partial w_k} = x_j \cdot \frac{\partial^{\pm} \varphi_j^H}{\partial w_k} + \varphi_j \cdot \frac{\partial x_j^H}{\partial w_k}$$

and the next corollary follows from Theorem 1.

Corollary 3

Under Assumption 1 it follows that

$$\frac{\partial \bar{X}_j}{\partial w_j} = \frac{\partial^+ \bar{X}_j^H}{\partial w_j} - x_j \cdot \frac{\partial \bar{X}_j}{\partial y}, \qquad \frac{\partial \bar{X}_j}{\partial w_j} = \frac{\partial^- \bar{X}_j^H}{\partial w_j} - \bar{X}_j \cdot \frac{\partial x_j}{\partial y},$$

$$\frac{\partial \bar{X}_k}{\partial w_j} = \frac{\partial^+ \bar{X}_k^H}{\partial w_j} \cdot \frac{\partial v_j / \partial y}{\partial v_k / \partial y} \quad \textit{and} \quad \frac{\partial \bar{X}_k}{\partial w_j} = \frac{\partial^- \bar{X}_k^H}{\partial w_j}$$

for $k \neq j.$

We observe that the relations in Corollary 3 differ in important ways from the corresponding (misleading) relations given in Small and Rosen (1981, pp. 116–18).

5. Special cases

Example 1: Urban travel demand

This example is taken from McFadden (1981, pp. 242–245). McFadden and his associates analyzed work-trip choice with four travel modes (bus, auto alone, rapid transit (BART), carpool) in the San Francisco Bay Area. One of the models they estimated was a multinomial logit model where the systematic part of the utility function of individual *i* was assumed to have the form

$$(5.1) \qquad v_j(w_j, y) = \alpha_j + \beta \frac{w_j}{y} + Z_{1j}\gamma_1 + Z_{2j}\gamma_2, \;\; \alpha_1 = 0$$

where y is the wage rate, w_j is the cost of alternative *j*, Z_{1j} is "on-vehicle time" of alternative *j*, Z_{2j} is "access time" of alternative *j*, and $\{\alpha_j\}$, β, γ_1, and γ_2 are unknown parameters. Train and McFadden (1978) have provided a theoretical justification of the utility as a function of the wage rate. Another justification is that if hours of work is given (such as full-time or part-time hours), then the wage rate is equal to the wage income, apart from a multiplicative constant.[2] The estimates of these parameters are given in Table 5.2, p. 244, in McFadden (1981). Below we give the uncompensated and compensated elasticities with respect to traveling costs and traveling times. Policy reforms that involve changes in traveling times may be implemented by reducing or increasing the number of transits and bus stops or transfers.

From (5.1) it follows that

$$(5.2) \quad \frac{\partial \varphi_j}{\partial y} = -\frac{\beta \varphi_j}{y^2}\left(w_j - \sum\nolimits_r w_r \varphi_r\right), \;\; \frac{\partial \varphi_j}{\partial w_j} = \frac{\beta}{y}\varphi_j(1-\varphi_j) \quad \text{and} \quad \frac{\partial \varphi_k}{\partial w_j} = -\frac{\beta \varphi_k \varphi_j}{y}.$$

Corollary 2 and (5.2) therefore imply the following compensated price elasticities:

$$\frac{\partial^+ \log \varphi_j^H}{\partial \log w_j} = \frac{\beta}{y}\sum\nolimits_{r \neq j} w_r \varphi_r, \qquad \frac{\partial^+ \log \varphi_k^H}{\partial \log w_j} = -\frac{\beta w_k \varphi_j}{y},$$

$$\frac{\partial^- \log \varphi_j^H}{\partial \log w_j} = \frac{\beta w_j}{y}(1-\varphi_j) \;\; \text{and} \;\; \frac{\partial^- \log \varphi_k^H}{\partial \log w_j} = -\frac{\beta w_j \varphi_j}{y}$$

for $k \neq j$. From the above results it follows that

---

[2] In most jobs, hours of work are fixed, possibly job-specific, and determined by institutional regulations or the nature of the jobs.

(5.3a) $$\frac{\partial^+ \log \varphi_j^H}{\partial \log w_j} - \frac{\partial^- \log \varphi_j^H}{\partial \log w_j} = \frac{\beta}{y}\sum_r (w_r - w_j)\varphi_r$$

and

(5.3b) $$\frac{\partial^+ \log \varphi_k^H}{\partial \log w_j} - \frac{\partial^- \log \varphi_k^H}{\partial \log w_j} = \frac{\beta(w_j - w_k)\varphi_j}{y}.$$

Similarly, it follows that the compensated elasticities with respect to traveling times are given by

$$\frac{\partial^+ \log \varphi_j^H}{\partial \log Z_{sj}} = \frac{\gamma_s Z_{sj}}{w_j}\sum\nolimits_{r\neq j} w_r \varphi_r, \quad \frac{\partial^+ \log \varphi_k^H}{\partial \log Z_{sj}} = -\frac{\gamma_s w_k Z_{sj}\varphi_j}{w_j},$$

$$\frac{\partial^- \log \varphi_j^H}{\partial \log Z_{sj}} = \gamma_s Z_{sj}(1-\varphi_j), \qquad \frac{\partial^- \log \varphi_k^H}{\partial \log Z_{sj}} = -\gamma_s Z_{sj}\varphi_j,$$

(5.4a) $$\frac{\partial^+ \log \varphi_j^H}{\partial \log Z_{sj}} - \frac{\partial^- \log \varphi_j^H}{\partial \log Z_{sj}} = \frac{\gamma_s}{w_j}\sum_{r\neq j} (w_r - w_j)\varphi_r$$

and

(5.4b) $$\frac{\partial^+ \log \varphi_k^H}{\partial \log Z_{sj}} - \frac{\partial^- \log \varphi_k^H}{\partial \log Z_{sj}} = \frac{\gamma_s Z_{sj}(w_j - w_k)\varphi_j}{w_j}.$$

From (5.3a, b) and (5.4a, b) we realize that the difference between the right and left marginal compensated elasticities may be substantial.

Example 2: Labor force participation with non-linear taxes

Consider the following model of labor force participation of married women. The women face the choice of working full-time (alternative 2) h with wage income w, or not working (alternative 1). Hours of work h is normalized to 1. Let a represent the disutility of fixed costs of working, y is non-labor income (husband's income) and $f(w, y)$ the function that transforms gross income labor income w and nonlabor income y to income after taxes. We assume that $f(w, y)$ is continuously differentiable. The utility of working is given by

(5.5a) $$U_2 = -a + \frac{\beta((f(w,y)^\alpha - 1)}{\alpha} + \varepsilon_2$$

and the utility of not working is given by

(5.5b) $$U_1 = \frac{\beta(f(0,y)^\alpha - 1)}{\alpha} + \varepsilon_1$$

where α and $\beta > 0$ are parameters and $\varepsilon_j, j = 1, 2,$ are random errors. When $\alpha = 0$ (5.5a, b) become

$$U_2 = -a + \beta \log f(w, y) \text{ and } \quad U_1 = \beta \log f(0, y) + \varepsilon_1.$$

Note that in this case utility is increasing in wage income. This implies that the right marginal compensated wage effects are equal to the corresponding uncompensated wage effects whereas the left marginal compensated wage effects differ from the corresponding uncompensated effects. Let φ_2 be the probability of participation. In the special case where the error terms are independent with standard Gumbel c.d.f. $\exp(-\exp(-x))$, then the probability of participation φ_2 (say) becomes

$$\varphi_2 = \frac{1}{1 + \exp(a + \beta(f(0, y)^\alpha - 1)/\alpha - \beta((f(w, y))^\alpha - 1)/\alpha)}.$$

Furthermore, it follows from (5.5a, b) and Corollary 1 that

$$\frac{\partial^+ \varphi_2^H}{\partial w} = \frac{\partial \varphi_2}{\partial w} = \beta(f(w, y))^{\alpha-1} f_1'(w, y) \varphi_2 (1 - \varphi_2),$$

$$\frac{\partial \varphi_2}{\partial y} = \beta(f(w, y)^{\alpha-1} f_2'(w, y) - f(0, y)^{\alpha-1} f_2'(0, y))(1 - \varphi_2)\varphi_2$$

and we therefore obtain that

$$\frac{\partial^- \varphi_2^H}{\partial w} = \frac{\partial \varphi_2}{\partial w} - \frac{\partial v_2(w, y)/\partial w}{\partial v_2(w, y)/\partial y} \cdot \frac{\partial \varphi_2}{\partial y} = \frac{\beta f(0, y)^{\alpha-1} f_2'(0, y)}{f_2'(w, y)} \varphi_2 (1 - \varphi_2).$$

Thus, in this case

(5.6)
$$\frac{\partial^- \varphi_2^H}{\partial w} - \frac{\partial^+ \varphi_2^H}{\partial w} = \frac{\partial v_2(w, y)/\partial w}{\partial v_2(w, y)/\partial y} \cdot \frac{\partial \varphi_2}{\partial y}$$

$$= \frac{\beta f_1'(w, y)(f(w, y)^{\alpha-1} f_2'(w, y) - f(0, y)^{\alpha-1} f_2'(0, y))}{f_2'(w, y)} (1 - \varphi_2)\varphi_2.$$

When $\alpha = 1$ the formula in (5.6) reduces to

(5.7)
$$\frac{\partial^- \varphi_2^H}{\partial w} - \frac{\partial^+ \varphi_2^H}{\partial w} = \frac{\beta f_1'(w, y)(f_2'(w, y) - f_2'(0, y))}{f_2'(w, y)} (1 - \varphi_2)\varphi_2.$$

The relation in (5.7) shows that even when utility is linear in disposable income the left and right marginal compensated wage effects are different if taxes are non-linear. If wife and husband are taxed separately, then $f_2'(w, y) - f_2'(0, y) = 0$, implying that the left and right marginal compensated wage effects are equal.

Example 3: The standard labor force participation model with non-linear taxes

Here we consider the labor force participation of married women when the model is assumed to be the standard textbook one where the worker is allowed to choose any level of continuous hours of work (subject to an upper limit on hours). Let h denote hours of work and w_2 the wage rate. We assume that the function $f(wh,y)$, which transforms labor income hw_2 and non-labor income y to income after taxes, is strictly concave in the wage rate w_2. In this case it follows that the woman would choose to work if $w_2 f_1'(0,y) > w_1$ and choose not to work otherwise, where w_1 is the woman's reservation wage (marginal rate of substitution at zero hours of work). In empirical applications it is necessary to represent the wage rate and the reservation wage by instrument variable equations. To this end assume that

$$\log w_2 = z_2 + \varepsilon_2 \quad \text{and} \quad \log w_1 = z_1 + \varepsilon_1$$

where z_1 and z_2 are the respective systematic parts of the log wage and log reservation wage equations that depend on suitable covariates, and ε_1 and ε_2 are zero mean random variables, possibly correlated, and independent of z_1 and z_2. Let F be the c.d.f. of $\varepsilon_1 - \varepsilon_2$. It thus follows that the probability of participation becomes

(5.8) $$\varphi_2 = P(z_2 + \log f_1'(0,y) + \varepsilon_2 > z_1 + \varepsilon_1) = F(\log f_1'(0,y) + z_2 - z_1).$$

From (5.8) it follows that

(5.9) $$\frac{\partial \varphi_2}{\partial z_2} = F'(\log f_1'(0,y) + z_2 - z_1)$$

and

(5.10) $$\frac{\partial \varphi_2}{\partial y} = F'(\log f_1'(0,y) + z_2 - z_1)\frac{f_{12}''(0,y)}{f_1'(0,y)}.$$

Hence, (5.9), (5.10) and Corollary 1 imply, with $v_1 = z_1$ and $v_2 = \log f_1'(0,y) + z_2$, that

$$\frac{\partial^+ \varphi_2^H}{\partial z_2} = \frac{\partial \varphi_2}{\partial z_2} \quad \text{and} \quad \frac{\partial^- \varphi_2^H}{\partial z_2} = 0.$$

Recall that the reservation wage does not depend on income. Thus, when the wage rate decreases the utility of the chosen alternative remains constant if the marginal wage rate at zero hours of work, $f_1'(0,y)w_2$, remains constant, which is obtained by a suitable increase of income. This means that neither the reservartion wage nor the marginal wage rate at zero hours of work changes, implying that the corresponding marginal effect is zero when $f''_{12}(0, y) \neq 0.$ If, however, $f''_{12}(0, y) = 0$, then the marginal compensated wage effect equals the corresponding uncompensated effect given in (5.9).

In the current Norwegian tax system wives and husbands are taxed separately, so $f''_{12}(0, y) = 0.$ In the previous Norwegian tax system, however, wives and husbands were taxed jointly when the wife's income was sufficiently low. Thus, in this latter case the function that transforms gross income to disposable income has the form $f(hw, y) = g(hw + y)$, which implies that $f''_{12}(0, y) = g''(y).$

6. Conclusions

In this paper we have discussed marginal compensated effects in discrete choice models and we have established the Slutsky equation for such models. As we have seen, the discrete Slutsky equation has the special feature that marginal compensated price effects in the case of a price increase differ from marginal compensated price effects in the case of a price decrease.

Funding: This study was funded by the Research Department of Statistics Norway and the Ragnar Frisch Centre of Economic Research, Oslo

References

Bhattacharya, D. (2015): Nonparametric welfare analysis for discrete choice. *Econometrica*, **83**, 617-649.

Bhattacharya, D. (2018): Empirical Welfare Analysis for Discrete Choice. *Quantitative Economics*, **9**, 571-615.

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 Random Utility Models? *Journal of Mathematical Psychology*, **39**, 90-98.

Dagsvik, J. K. and A. Karlström (2005): Compensating Variation and Hicksian Choice Probabilities in Random Utility Models that are Nonlinear in Income. *Review of Economic Studies*, **72**, 57-76.

Dagsvik, J. K. and S. Strøm (2006): Sectoral Labor Supply, Choice Restrictions and Functional Form. *Journal of Applied Econometrics*, **21**, 803-826.

Dagsvik, J. K., M. Locatelli and S. Strøm (2009): Tax Reform, Sector Specific Labor Supply and Welfare Effects. *Scandinavian Journal of Economics,* **111**, 265-287.

Dagsvik, J. K., S. Strøm and M. Locatelli (2019): Compensated Marginal Effects and the Slutsky Equation in Discrete Labor Supply Models. CESifo Working Paper, No.7493.

Dubin, J., and D. McFadden (1984): An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, **52**, 345-362.

Hanemann, W. M. (1984): Discrete/Continuous Choice of Consumer Demand. *Econometrica*, **52**, 541-561.

Hausman, J. A. (1981): Exact Consumer's Surplus and Deadweight Loss. *American Economic Review*, **71**, 662-676.

Hicks, J. R. (1936): *Value and Capital*, Oxford University Press.

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

Kornstad, T. and T. O. Thoresen (2006): Effect of Family Policy Reform in Norway. *Fiscal Studies*, **27**, 339-371.

McFadden, D. (1978): Modelling the Choice of Residential Location. In A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (eds.): *Spatial Interaction Theory and Planning Models*. North Holland, Amsterdam.

McFadden, D. (1981): Econometric models of probabilistic choice. In C. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Massachusetts.

Niemeier, D. A. (1997): Accessibility: An Evaluation Using Consumer Welfare. *Transportation*, **24**, 377–396.

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 Honor of John S. Chipman*. Routledge, London.

Resnick. S. I. (1987): *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag, Berlin.

Small, K. A. and H. S. Rosen (1981): Applied Welfare Analysis in Discrete Choice models. *Econometrica*, **49**, 105-130.

Slutsky, E. E. (1915): Sulla Teoria del Bilancio del Consumatore. *Giornale degli Economisti*, **51**, 1–26.

Sneddon, I. N. (1957): *Elements of Partial Differential Equations*. McGraw Hill, New York.

Train, K. And D. McFadden (1978): The Goods-Leisure Tradeoff and Disaggregate Work Trip Mode Choice Models. *Transporation Research*, **12**, 349-353.

Van Soest, A. (1995): A Structural Model of Family Labor Supply: A Discrete Choice Approach. *Journal of Human Resources*, **30**, 63-88.

Varian, H. (1992): *Microeconomic Analysis* (3rd ed.). W. W. Norton, New York.

Appendix

Lemma 1

Let $U_j(x,y)=v_j(x,y)+\varepsilon_j,\ j=1,2,...,m,$ *with random error terms* $\{\varepsilon_j\}$ *that have joint c.d.f. F with deterministic terms* $\{v_j(w,y)\}$ *that are strictly increasing in y. For simplicity, write* $v_j=v_j(w_j,y)$ *for the ex ante systematic part of the utility function and let* $\tilde{w}$ *be the vector of ex post prices. Then*

(i) $$P(\max_{r\le m} U_r(w_r,y)=U_j(w_j,y)=\max_{r\le m} U_r(\tilde{w}_r,Y)=U_k(\tilde{w}_k,Y)\in du, Y\in dz)$$
$$=F''_{jk}(u-\psi_1(z),...,u-v_j,...,u-v_k(w_2,z),u-\psi_m(z))v_k(w_k,dz)du$$

for $y_j>y\ge y_k,\ k\ne j,\ k,j>0,$ *and*

(ii) $$P(\max_r U_r(w_r,y)=U_j(w_j,y)=\max_r U_r(\tilde{w}_r,Y)\in du, Y=y_j)$$
$$=F'_j(u-\psi_1(y_j),...,u-v_j,...,u-\psi_m(y_j))du$$

where $\psi_j(z)=\max(v_j,v_j(\tilde{w}_j,z))$ *and* y_j *is determined by* $v_j=v_j(\tilde{w}_j,y_j)$.

Proof of Lemma 1:

Consider first the proof of (i). Let J and $\tilde{J}$ denote the ex-ante and ex-post choice given that the ex-ante and ex-post utility levels of the chosen alternatives are equal. Let F be the joint c.d.f. of $(\varepsilon_1,\varepsilon_2,...,\varepsilon_m)$. For notational convenience, let $U_r=U_r(w_r,y)$ and $\tilde{U}_r(Y)=U_r(\tilde{w}_r,Y)$. We have that

$$\{J=1,\tilde{J}=2\}\Leftrightarrow\{\max_{r\ne\{1,2\}} U_r\le U_1, \max_{r\ne 2}\tilde{U}_r(Y)\le\tilde{U}_2(Y)=U_1\}.$$

For alternative 1 to be the most preferred alternative ex-ante and alternative 2 the most preferred alternative ex-post, one must have $\tilde{U}_2(Y)=U_1>U_2$. Hence, we must have that $Y>y_2$. Furthermore, since alternative 2 is the most preferred one ex-post, $\tilde{U}_2(Y)>\tilde{U}_1(Y)$, which implies that $\tilde{U}_1(Y)<U_1$ and $Y<y_1$. Accordingly, the event $\{J=1,\tilde{J}=2\}$ has positive probability only if $y_1>y_2$. Moreover, the event $\{\tilde{U}_2(z)=U_1\}$ implies that $\{\tilde{U}_1(z)\le U_1, U_2\le U_1\}$. Accordingly, the relation above yields

$$\{J=1,\tilde{J}=2,Y=z\}\Leftrightarrow\{\max_{r\ne\{1,2\}}\max(U_r,\tilde{U}_r(z))\le U_1=\tilde{U}_2(z)\}.$$

Thus, the corresponding probabilities are therefore given by

$$P(J=1,\tilde{J}=k,Y\in[z,z+\Delta z),U_1\in(u,u+\Delta u))$$
$$=P(\max_{r\ne 1} U_r\le u, \max_{r\ne 2}\tilde{U}_r(Y)\le\tilde{U}_2(Y), Y\in(z,z+\Delta z), U_1\in(u,u+\Delta u))+o(\Delta u)$$
$$=P(\max_{r\ne 1} U_r\le u, \max_{r\ne 2}\tilde{U}_r(z)\le u, \tilde{U}_2(z)\le u<\tilde{U}_2(z+\Delta z), U_1\in(u,u+\Delta u))+o(\Delta u)$$
$$=P(\max_{r\ne\{1,2\}}\max(U_r,\tilde{U}_r(z))\le u, \tilde{U}_2(z)\le u<\tilde{U}_2(z+\Delta z), U_1\in(u,u+\Delta u))+o(\Delta u)$$
$$=P(\max_{r\ne\{1,2\}}(\psi_r(z)+\varepsilon_r)\le u, \tilde{U}_2(z)\le u<\tilde{U}_2(z+\Delta z), U_1\in(u,u+\Delta u))+o(\Delta u)$$

$$= F_1'(u - v_1, u - v_2(\tilde{w}_2, z), u - \psi_3(z), ..., u - \psi_m(z))\Delta u$$

$$- F_1'(u - v_1, u - v_2(\tilde{w}_2, z + \Delta z), u - \psi_3(z), ..., u - \psi_m(z))\Delta u + o(\Delta u)$$

$$= F_{12}''(u - v_1, u - v_2(\tilde{w}_2, y), u - \psi_3(z), ..., u - \psi_m(z))\Delta v_2(z)\Delta u + o(\Delta z)\Delta u.$$

This proves (i) of Lemma A1. Consider now the second part (ii). We have that

$$P(\max_{r \neq j} U_r \leq U_j \geq \max_{r \neq j} U_r(\tilde{w}_r, Y) \in du, Y = y_j)$$

$$= P(\max_{r \neq j} U_r \leq U_j \geq \max_{r \neq j} U_r(\tilde{w}_r, y_j) \in du)$$

$$= P(\max_{r \neq j}(\varepsilon_r + \psi_r(y_j)) \leq v_j + \varepsilon_j \in du)$$

$$= F_j'(u - \psi_1(y_j), ..., u - v_j, ..., u - \psi_m(y_j))du,$$

which proves the second part.

Q.E.D.

Lemma 2

Let $b_1, b_2, ..., b_m,$ *be positive constants and* $F(x_1, x_2, ..., x_m)$ *a multivariate c.d.f. defined on* R^m *that possesses a p.d.f. Assume that F satisfies the partial differential equation*

$$\frac{\partial F(u + x_1, u + x_2, ..., u + x_m)b_j}{\partial x_j} = P_j(x_1, x_2, ..., x_m)\sum_k \frac{\partial F(u + x_1, u + x_2, ..., u + x_m)b_k}{\partial x_k}$$

for some positive function $P_j(x_1, x_2, ..., x_m) \in (0,1),$ *defined on* R^m. *Then F must have the form*

$$F(x_1, x_2, ..., x_m) = \zeta(G(e^{-b_1 x_1}, e^{-b_2 x_2}, ..., e^{-b_m x_m}))$$

where $\zeta : (0, \infty) \to [0,1]$ *is a strictly decreasing function and* $G(z_1, z_2, ..., z_m) : [0, \infty)^m \to [0, \infty)$ *is a strictly decreasing and linear homogenous function. In other words, F is a multivariate extreme value c.d.f. (Resnick, 1987).*

Proof of Lemma 2:

Assume first that $b_j = 1$ for all *j*. By applying the method of Lagrange for solving partial differential equations (Sneddon, 1957), the first stage is to solve the differentical equation

$$\frac{du}{dx_j} = P_j$$

which yields

(A.1) $\quad u - Q_j(x_1, x_2, ..., x_m) = C_j$

for all *j*, where $\partial Q_j / \partial x_j = P_j$ and C_j is a constant. The equation in (A.1) implies that

$$Q_j(x_1,x_2,...,x_m)=Q_1(x_1,x_2,...,x_m)+C_j-C_1=Q_1(x_1,x_2,...,x_m)+c_j.$$

Then it follows that the solution of the differentical equation in Lemma 2 in this case must satisfy $\xi_j(u-Q_1(x_1,x_2,...,x_m),F(u+x_1,u+x_2,...,u+x_m))=0$ for some arbitrary continuously differentiable function ξ in two variables. The latter equation implies that

(A.2) $$F(u+x_1,u+x_2,...,u+x_m)=\zeta_j(e^{-u}\exp Q_1(x_1,x_2,...,x_m)).$$

for some suitable positive and continuously differential function $\zeta_j\in[0,1]$. Evidently, ζ_j must be strictly decreasing and independent of *j* because the left hand side of (A.2) is a c.d.f. that is independent of *j*. With no loss of generality we can rewrite (A.2) as

(A.3) $$F(u+x_1,u+x_2,...,u+x_m)=\zeta(e^{-u}G(e^{-x_1},e^{-x_2},...,e^{-x_m}))$$

where $G(z_1,z_2,...,z_m)=\exp Q_1(-\log z_1,-\log z_2,...,\log z_m)$. Eq. (A.3) implies that

$$\zeta(G(e^{-u-x_1},e^{-u-x_2},...,e^{-u-x_m}))=\zeta(e^{-u}G(e^{-x_1},e^{-x_2},...,e^{-x_m}))$$

which yields

$$G(e^{-u-x_1},e^{-u-x_2},...,e^{-u-x_m})=e^{-u}G(e^{-x_1},e^{-x_2},...,e^{-x_m})$$

and thus establishes linear homogeneity of *G*. Furthermore, we realize that in the general case the partial differential equation in Lemma 2 with $b_j\neq 1$ must have the solution

$$F(x_1,x_2,...,x_m)=\zeta(G(e^{-b_1x_1},e^{-b_2x_2},...,e^{-b_mx_m})).$$

Q.E.D.

Proof of Proposition 1:

Note first that

$$e_j(w_j,u)>y \iff v_j(w_j,y)+\varepsilon_j<u.$$

Let $y_k, k=1,2,...,m,$ be positive real numbers. Then the relation above implies that

$$P(e_1(w_1,u)>y_1,e_2(w_2,u)>y_2,...,e_m(w_m,u)>y_m)$$
$$=P(v_1(w_1,y_1)+\varepsilon_1<u,v_2(w_2,y_2)+\varepsilon_2<u,...,v_m(w_m,y_m)+\varepsilon_m<u).$$

Accordingly,

(A.4) $$P(e(w,u)>y)=P(\max_{k\le m}(v_k(w_k,y)+\varepsilon_k)<u)$$
$$=F(u-v_1(w_1,y),u-v_2(w_2,y),...,u-v_M(w_M,y))$$

and

(A.5) $$P(J(w,y)=j,e(w,u)\in dy)=P(\min_{k\le m}e_k(w_k,u)=e_j(w_j,u)\in dy)$$

$$= F_j(u - v_1(w_1,y), u - v_2(w_2,y), ..., u - v_m(w_m,y))v'_{j2}(w_j,y)dy.$$

From (A.4) and (A.5) it follows that

(A.6) $$P(J(w,y) = j \mid e(w,u) = y) = \frac{F_j'(u - v_1(w_1,y), u - v_2(w_2,y), ..., u - v_M(w_m,y))v'_{j2}(w_j,y)}{\sum_{r=1}^{m} F_r'(u - v_1(w_1,y), u - v_2(w_2,y), ..., u - v_M(w_m,y))v'_{r2}(w_j,y)}.$$

Since the choice probability $P(J(w,y) = j)$ is independent of u it follows from Lemma 2 with $b_j = v'_{j2}(w_j,y)$ that F must have the form

$$F(x_1, x_2, ..., x_m) = \zeta(G(e^{-b_1x_1}, e^{-b_2x_2}, ..., e^{-b_mx_m})).$$

But since F is the joint c.d.f. of the vector of random error terms Assumption 1 implies that $b_j = v'_{j2}(w_j,y)$ must be a constant, independent of j.

Q.E.D.

Let y_j be determined by $v_j(w_j,y) = v_j(\tilde{w}_j, y_j)$. That is, y_j is the ex-post income that ensures that the deterministic parts of the ex-ante utility and ex-post utility of alternative *j* are equal. If alternative *j* belongs to the ex ante choice set but not the ex post choice set, we define $y_j = \infty$. If alternative *j* belongs to the ex post choice set but not the ex ante choice set, we define $y_j = 0$.

Lemma 3

Under Assumption 1 the Hicksian (compensated) choice probability of changing from alternative j to alternative k is given by[3]

(A.7) $$Q^H(j,k) = -\int_{y_k}^{y_j} H''_{jk}(\psi_1(z), \psi_2(z), ..., \psi_m(z))\tilde{v}_k(\tilde{w}_k, dz)$$

when $k \neq j$, $j, k > 0$, $y_j > y_k$, *where* $\psi_r(z) = \max(v_r(w_r,y), \tilde{v}_r(\tilde{w}_r, z))$ *for all* r. *Furthermore, when* $j = k$, *then*

(A.8) $$Q^H(j,j) = H'_j(\psi_1(y_j), \psi_2(y_j), ..., \psi_m(y_j)).$$

When $y_j \leq y_k$, $j \neq k$, $j, k > 0$, *then* $Q^H(j,k) = 0$.

[3] There is an error in the corresponding formula for $Q^H(j,j)$ in Theorem 4 in Dagsvik and Karlström (2005).

Lemma 3 implies that $Q^H(j,k)=0$ when $\tilde{w}_j < w_j, \tilde{w}_r = w_r$ for $r \neq j,k,$ $\tilde{w}_k \geq w_k$. Suppose that $v_r(w_r, y) = v_r^*(y - w_r)$ for all r where v_r^* is an increasing function. Then, under the assumptions of Theorem 1 it follows that for $j \neq k$,

$$Q^H(j,k) = \int_{y_k}^{y_j} \frac{\partial H_j'(\psi_1(z),\psi_2(z),...,\psi_m(z))dz}{\partial w_k} \quad \text{and } Q^H(j,j) = \varphi_j(w,y)$$

when $y_j \geq y_k$. Thus, we realize that the integrand of the integral in the expression for $Q^H(j,k)$ in this case can be obtained from the Marshallian choice probabilities as functions of the deterministic functions $\{\psi_r(z)\}$.

Lemma 3 has originally been derived by Dagsvik and Karlström (2005). In this paper we provide a simplified proof given in the appendix. Note that $Q^H(j,j)$ given in (A.8) has the same structure as a choice probability, namely the probability of choosing alternative *j* when the utility of alternative *j* equals $\psi_j(z)+\varepsilon_j$, j = 1, 2, …, *m*.

. .

Proof of Lemma 3:

From Lemma 1(i) (with the same notation as in Lemma 1) it follows that

(A.9) $P(J=j, \tilde{J}=k, Y \in [z, z+\Delta z))$

$$= v_{k2}'(\tilde{w}_k, z)\Delta z \int_{-\infty}^{\infty} F_{jk}''(u-\psi_1(z),..,u-v_k(\tilde{w}_k,z),u-\psi_3(z),...,u-\psi_m(z))du + o(\Delta z).$$

Furthermore, we know that

$$H_j'(v_1,v_2,...,v_m) \equiv P(v_j+\varepsilon_j = \max_r(v_r+\varepsilon_r)) = \int_{-\infty}^{\infty} F_j'(x-v_1,x-v_2,..,x-v_m)dx.$$

Evidently, differentiation under the integral above is allowed in this case, which yields that

$$-v_{k2}'(\tilde{w}_k,z)\Delta z H_{jk}''(\psi_1(z),v_k(\tilde{w}_2,z),\psi_3(z),..,\psi_m(z))$$

$$= v_{k2}'(\tilde{w}_k,z)\Delta z \int_{-\infty}^{\infty} F_{jk}''(x-\psi_1(z),x-v_k(\tilde{w}_k,z),..,x-\psi_m(z))dx$$

$$= P(J=1, \tilde{J}=k, Y \in [z,z+\Delta z)) + o(\Delta z).$$

Furthermore, since alternative 1 is chosen ex ante and alternative 2 ex post it must be the case that $U_j(w_j,y) > U_j(\tilde{w}_j,Y)$ and $U_k(\tilde{w}_k,Y) > U_k(w_k,y)$ implying that $v_j(w_j,y) > v_j(\tilde{w}_j,Y)$ and $v_k(\tilde{w}_k,Y) > v_k(w_k,y)$. Hence, it must be true that the probability in (A.9) wanished unless

$y_k \leq Y \leq y_j$. By integrating (A.9) with respect to z between y_k and y_j yields (A.7). The relation in (A.8) follows from Lemma 1 (ii).

Q.E.D.

Proof of Theorem 1:

By assumption $v_r(w_r, y)$ is strictly decreasing in prices and strictly increasing in income. Let the price of alternative *j* increase from w_j to $\tilde{w}_j = w_j + \Delta w_j$ where Δw_j is small and positive. Then, $y_j(\tilde{w}_j) > y$ and $y_r = y$ for $r \neq j$. Hence, it follows from Lemma 2 that $Q^H(r, j) = 0$ for $r \neq j$. Furthermore, from the definition of $y_j(u)$ it follows that $v_j(u, y_j(u)) = v_j(w_j, y)$ for any given real *u*. By implicit differentiation of the latter equation with respect to *u* yields

(A.10) $$\frac{\partial y_j(w_j)}{\partial w_j} = \frac{\partial y_j}{\partial w_j} = -\frac{v'_{j1}(w_j, y)}{v'_{j2}(w_j, y)} = -\frac{v'_{j1}}{v'_{j2}}.$$

where $y_j = y_j(w_j)$. Since $y_j > y$, $\max_r(v_r(w_r, y), v_r(w_r, y_j)) = v_r(w_r, y_j)$ and we get from Lemma 2 and (A.10) that

(A.11) $$\sum_r Q^H(r, j) - \varphi_j = Q^H(j, j) - \varphi_j = H'_j(v_1(w_1, y_j), v_2(w_2, y_j), ..., v_j(w_j, y), .., v_m(w_m, y_j))$$

$$-H'_j(v_1(w_1, y), v_2(w_2, y), ..., v_m(w_m, y))$$

$$= \frac{\partial y_j(w_j)}{\partial w_j} \sum_{r \neq j} H''_{jr}(v_1(w_1, y), v_2(w_2, y), ...v_j(w_j, y), .., v_m(w_m, y))v'_{r2}(w_r, y)\Delta w_j + o(\Delta w_j)$$

$$= \left(H''_{jj}(v_1(w_1, y), v_2(w_2, y), ..., v_m(w_m, y)) - \frac{\partial H'_j(v_1(w_1, y), v_2(w_2, y), ..., v_m(w_m, y))}{\partial y} \right) \frac{v'_{j1}(w_j, y)}{v'_{j2}(w_j, y)} \Delta w_j + o(\Delta w_j)$$

$$= \frac{\partial H'_j(v_1(w_1, y), v_2(w_2, y), ..., v_m(w_m, y))}{\partial w_j} - \frac{\partial H'_j(v_1(w_1, y), v_2(w_2, y), ..., v_m(w_m, y))}{\partial y} \cdot \frac{v'_{j1}(w_j, y)}{v'_{j2}(w_j, y)} \Delta w_j + o(\Delta w_j)$$

$$= \frac{\partial \varphi_j}{\partial w_j} + \frac{\partial y_j}{\partial w_j} \cdot \frac{\partial \varphi_j}{\partial y} = \frac{\partial \varphi_j}{\partial w_j} - \frac{v'_{j1}}{v'_{j2}} \cdot \frac{\partial \varphi_j}{\partial y}.$$

Since

$$\frac{\partial^+ \varphi_j^H}{\partial w_j} = \lim_{\Delta w_j \to 0} \left(\frac{\sum_r Q^H(r, j) - \varphi_j}{\Delta w_j} \right)$$

the first part of the theorem follows from (A.11).

Consider next the corresponding cross price effects. That is, we consider the marginal compensated effect on φ_k when $k \neq j$. We have that $Q^H(r, k) = 0$ for $r \neq j$ and $Q^H(j, k) > 0$. From Lemma 2 we obtain that

(A.12) $$Q^H(j,k) = -\int_y^{y_j(\tilde{w}_j)} H''_{jk}(v_1(w_1,x),..,v_j(w_j,y),..,v_k(w_k,x),..,v_m(\tilde{w}_m,x))v'_{k2}(w_k,x)dx$$

which together with (A.10) imply that

(A.13) $$Q^H(j,k) = -H''_{jk}(v_1(w_1,y),...,v_m(w_m,y))v'_{k2}(w_k,y)(y_j(\tilde{w}_j)-y)+o(\Delta w_j)$$

$$= -\frac{\partial H'_k(v_1(w_1,y),...,v_m(w_m,y))v'_{k2}(w_k,y)(y_j(\tilde{w}_j)-y)}{v'_{j1}(w_k,y)\partial w_j}+o(\Delta w_j)$$

$$= -\frac{\partial \varphi_k}{\partial w_j}\cdot\frac{v'_{k2}(w_k,y)}{v'_{j1}(w_k,y)}\cdot\frac{\partial y_j(w_j)\Delta w_j}{\partial w_j}+o(\Delta w_j) = \frac{\partial \varphi_k}{\partial w_j}\cdot\frac{v'_{k2}(w_k,y)\Delta w_j}{v'_{j2}(w_k,y)}+o(\Delta w_j).$$

Furthermore, it follows that

(A.14) $$Q^H(k,k) = \varphi_k.$$

From (A.12) and (A.13) it thus follows that

$$\frac{\partial^+ \varphi_k^H}{\partial w_j} = \lim_{\Delta w_j\to 0}\left(\frac{Q^H(j,k)+Q^H(k,k)-\varphi_k}{\Delta w_j}\right) = \lim_{\Delta w_j\to 0}\frac{Q^H(j,k)}{\Delta w_j} = \frac{\partial \varphi_k}{\partial w_j}\cdot\frac{v'_{k2}(w_k,y)}{v'_{j2}(w_k,y)}.$$

Consider next the case where Δw_j is negative. Then it follows that $y_j < y$ so that $\psi_r(y_j) = v_r(w_r,y)$ implying that $Q^H(j,j)=\varphi_j$. From Lemma 3 we get that

$$Q^H(r,j) = -H''_{rj}(v_1(w_1,y),..,v_j(w_j,y),..,v_k(w_k,y),..,v_m(w_m,y))v'_{j2}(\tilde{w}_j,x)(y-y_j(\tilde{w}_j))+o(\Delta w_j)$$

$$= H''_{rj}(v_1(w_1,y),..,v_j(w_j,y),..,v_m(w_m,y))v'_{j2}(w_j,y)\Delta w_j y'_j(w_j)+o(\Delta w_j)$$

$$= -H''_{kj}(v_1(w_1,y),..,v_j(w_j,y),..,v_m(w_m,y))\cdot\frac{v'_{j2}(w_j,y)\Delta w_j v'_{j1}(w_j,y)}{v'_{j2}(w_j,y)}+o(\Delta w_j)$$

$$= -H''_{kj}(v_1(w_1,x),..,v_j(w_j,y),..,v_m(w_m,y))\Delta w_j v'_{j1}(\tilde{w}_j,y)+o(\Delta w_j).$$

Consequently, we obtain that

$$\sum_r Q^H(r,j)-\varphi_j = \sum_{r\neq j} Q^H(r,j)$$

$$= -\sum_{r\neq j} H''_{rj}(v_1(w_1,x),..,v_j(w_j,y),..,v_m(w_m,y))\Delta w_j v'_{j1}(\tilde{w}_j,y)+o(\Delta w_j)$$

$$= -\sum_{k\neq j} H''_{kj}(v_1(w_1,x),..,v_j(w_j,y),..,v_m(w_m,y))\Delta w_j v'_{j1}(w_j,y)+o(\Delta w_j)$$

$$= -\sum_{k\neq j}\frac{\partial \varphi_k}{\partial w_j}\cdot\Delta w_j + o(\Delta w_j) = \frac{\partial \varphi_j}{\partial w_j}\Delta w_j + o(\Delta w_j)$$

which implies that

$$\frac{\partial \varphi_j^H}{\partial w_j} = \frac{\partial \varphi_j}{\partial w_j}.$$

We also have that $Q^H(j,k) = 0$ when $k \neq j,$ and that

$$Q^H(k,k) = H_k'(v_1(w_1,y),..,v_{j-1}(w_{j-1},y),v_j(\tilde{w}_j,y),v_{j+1}(w_{j+1},y),...,v_m(w_m,y)).$$

Therefore, we get that

$$\sum_r Q^H(r,k) - \varphi_k = Q^H(k,k) - \varphi_k$$

$$= H_k'(v_1(w_1,y),..,v_j(\tilde{w}_j,y),...,v_m(w_m,y)) - \varphi_k.$$

By first order Taylor expansion the last expression becomes

$$H_{kj}''(v_1(w_1,y),v_2(w_2,y),...,v_m(w_m,y))v_{j1}'(w_j,y)\Delta w_j + o(\Delta w_j) = \frac{\partial \varphi_k \Delta w_j}{\partial w_j} + o(\Delta w_j)$$

which implies that

$$\frac{\partial \varphi_k^H}{\partial w_j} = \frac{\partial \varphi_k}{\partial w_j}.$$

Q.E.D.