APPLIED WELFARE ECONOMICS WITH DISCRETE CHOICE MODELS: IMPLICATIONS FOR MODEL SPECIFICATION

Richard Batley\textsuperscript{a} & J Nicolás Ibáñez\textsuperscript{a,b}

\textsuperscript{a} Institute for Transport Studies, University of Leeds, UK
\textsuperscript{b} Institute for Prospective Technological Studies, DG Joint Research Centre, European Commission, Sevilla, Spain

ABSTRACT

Small & Rosen’s 1981 paper has played an influential role in promoting the application of discrete choice models to the welfare analysis of public policy interventions. The contribution of the present paper is to review the theoretical basis of Small & Rosen, with a view to strengthening practical advice. We demonstrate that Small & Rosen’s welfare measure implies four requirements on the specification of deterministic utility within the discrete choice model, namely: (i) for each alternative, equivalence (in absolute terms) between the conditional marginal utilities of income and price; (ii) common conditional marginal utility of income across alternatives; (iii) common conditional marginal utility of price across alternatives; and (iv) independence of the conditional marginal utility of income from prices. We show that a discrete choice model specified in this manner yields a probabilistic demand function that observes the fundamental properties of adding-up, negativity, homogeneity and symmetry, and implies a null income effect. We conclude that Small & Rosen’s welfare measure is valid only if the discrete choice model complies with the four specification requirements, and income effects are not relevant to the context of application.

(185 words)
WORDCOUNT: 10,275 words

PAGECOUNT: 51

JEL CLASSIFICATION: B41 Economic Methodology; D01 Microeconomic Behaviour, Underlying Principles

KEYWORDS: Discrete choice models, welfare, log sum, homogeneity, symmetry

AUTHOR FOR CORRESPONDENCE:
Dr Richard Batley
Principal Research Fellow & Leader of the Economics Research Group
Institute for Transport Studies
University of Leeds
Leeds LS2 9JT
United Kingdom
Telephone: +44 (0) 113 343 1789
Fax: +44 (0) 113 343 5334
E-mail: R.P.Batley@its.leeds.ac.uk
1 INTRODUCTION

The apparatus of the Random Utility Model (RUM) modelling first emerged in the early 1960s, with Marschak (1960) and Block & Marschak (1960) translating models originally developed for discriminant analysis in psychophysics (Thurstone, 1927) to the alternative domain of discrete choice analysis in economics. Whilst some researchers were quick to see its practical potential (e.g. McFadden, 1968, 1975), it was not until the late 1970s and early 1980s that RUM was equipped with a reasonably comprehensive theoretical rationale in terms of the economics of consumption. An important tenet of this rationale was the link between discrete choice and welfare, which established a basis for using RUM to inform public policy, and paved the way for the plethora of applications which have been witnessed over the last 30 years.

It will be helpful to clarify precisely what we mean by ‘discrete choice’, since Small & Rosen (1981) - which will be referred to simply as ‘S&R’ in the remainder of this paper - suggest three alternative rationales, as follows. First, commodities may be available in continuous quantities but in a limited number of varieties. Second, goods may be supplied in discrete units of such magnitude that only a small number of those units are typically consumed (in this case, S&R cite the example of travel mode choice). Third, if the search for the optimal consumption bundle entails a choice between alternative corner solutions, then the problem is reduced to discrete units. The first rationale is the focus of S&R’s paper, and will also therefore be the focus of the present paper. Against this rationale, S&R introduce a general model of demand comprising both continuous and discrete components. That is to say, an individual is represented as choosing a quantity of a continuous commodity conditional upon discrete choice.

Whilst not overlooking the notable contributions of McFadden (1981) and Williams (1977), S&R's analysis has proved particularly influential in establishing a basis
upon which discrete choice models can be applied to welfare economics. S&R state that ‘The purpose of [their] paper is to demonstrate that the conventional methods of applied welfare economics can be generalised to handle cases in which discrete choices are involved’ (S&R, 1981, p106). Furthermore, they remark that: ‘Throughout, the emphasis is on providing rigorous guidelines for carrying out empirical work’ (p106). It is notable that, despite the intensity with which RUM has been applied over the last 30 years, S&R’s paper has stood the test of time; the key propositions of the paper remain largely unchallenged and continue to underpin the analysis of significant public policy interventions.

That said, in the years following its publication, a small but significant literature (e.g. Hau, 1985, 1987; Jara-Díaz & Farah, 1988; Jara-Díaz, 1990; Karlström, 1999; Karlström & Morey, 2004) has exposed various assumptions underpinning S&R and, in so doing, clarified the properties of resulting measures of welfare. In particular, these contributors have clarified the Marshallian basis of S&R, and stimulated research effort towards developing Hicksian analogues. The present paper will seek to contribute to the aforementioned literature, by furthering understanding of S&R, especially in a manner that appeals to S&R’s aspiration to provide ‘...rigorous guidelines for carrying out empirical work’. More specifically, our paper will offer five substantive contributions.

First, having outlined the demand problem considered by S&R in full and explicit terms, the paper will articulate the concept of a probabilistic demand function, and expose the assumptions underlying that concept.

Second, the paper will consider the manner by which the Slutsky equation applies to the discrete and continuous components of demand, from both individual-level and aggregate perspectives. In this regard, the present paper will expose the assumptions underpinning S&R’s derivation of the Slutsky equation.

The third contribution of the paper will be to reconcile S&R’s model of discrete-continuous demand with four fundamental properties of demand functions,
namely ‘adding up’, ‘negativity’, ‘homogeneity’ and ‘symmetry’. For a restricted case of S&R’s model involving only the probabilistic demand, the paper will identify particular requirements on model specification, in order that the aforementioned properties hold.

Fourth, the paper will review the rationale followed by S&R in deriving welfare measures from discrete choice models, and reconcile this rationale with the aforementioned properties of demand functions. It will be shown that the ‘log sum’ measure of consumer surplus implies very particular requirements on model specification, consistent with those relevant to the properties of demand functions.

2 A MODEL OF DISCRETE-CONTINUOUS DEMAND

2.1 Introduction

This section will summarise in formal terms S&R’s model of discrete-continuous demand. Having considered individual-level demand, which (strictly speaking) is the relevant perspective for the aforementioned properties of demand functions, the paper will proceed to consider the implications for aggregate demand, which is the perspective more commonly adopted by discrete choice analysts. That is to say, analysts usually interpret the probabilistic demand function associated with discrete choice as deriving from inter-individual rather than intra-individual variation in preferences.
2.2 Discrete-continuous demand, in principle

S&R consider the problem of an individual consumer choosing quantities of three goods \((x_1, x_2, x_n)\), where the first two goods are mutually exclusive, and the third good acts as the numeraire. More formally, the individual faces the following maximisation problem:

\[
\begin{align*}
\text{Max} & \quad u = u(x_1, x_2, x_n) \\
\text{s.t.} & \quad p_1 x_1 + p_2 x_2 + x_n = y \\
& \quad x_1 x_2 = 0 \\
& \quad x_1, x_2, x_n \geq 0
\end{align*}
\]

where \(x\) is quantity, \(p\) is price and \(y\) is income.

The constraint \(x_1 x_2 = 0\), which precludes positive consumption of both goods 1 and 2, is especially pertinent to the present paper, since it embodies S&R’s notion of discrete-continuous demand. The notion of discrete-continuous demand entails a two-stage consumption decision. The individual first chooses between goods 1 and 2 according to which yields the greater utility. Let us denote this:

\[
\hat{u} = \hat{v}_k(p, y) = \text{Max}\{\hat{v}_1(p, y), \hat{v}_2(p, y)\}
\]

where \(\hat{v}\) is conditional indirect utility, \(y\) is the income required to achieve the maximum direct utility \(\hat{u}\) both unconditionally and conditionally, and \(k\) indexes the chosen (i.e. utility maximising) good, i.e. \(k = 1\) if \(\hat{v}_1 \geq \hat{v}_2\), or \(k = 2\) otherwise\(^1\).

Having chosen between goods 1 and 2, the individual then selects a positive continuous quantity of only the chosen good.

More specifically, in solving (1), S&R adopt a particular interpretation of Roy’s identity, which yields conditional uncompensated demands for goods 1 and 2 as follows:

\(^1\) Furthermore, let us arbitrarily assign \(k = 1\) to the case \(\hat{v}_1 = \hat{v}_2\), so as to avoid ambiguities in the event of ties.
\[- \frac{\partial v}{\partial p_1} = \begin{cases} -\frac{\partial \tilde{v}_1}{\partial p_1} = \bar{x}_1 & \text{if } k = 1 \\ -\frac{\partial \tilde{v}_2}{\partial p_1} = 0 & \text{if } k = 2 \end{cases}\]

\[- \frac{\partial v}{\partial p_2} = \begin{cases} -\frac{\partial \tilde{v}_1}{\partial p_2} = 0 & \text{if } k = 1 \\ -\frac{\partial \tilde{v}_2}{\partial p_2} = \bar{x}_2 & \text{if } k = 2 \end{cases}\]

Although (2) and (3) usefully introduce the ‘primal’ problem of utility maximisation, S&R develop most of their analysis (an important exception being their discussion of welfare, reviewed in section 6 of the present paper) from the ‘dual’ perspective of minimising the expenditure necessary to achieve the utility $u^*$:

$$e(p_1, p_2, u^*) = \bar{e}_k(p_k, u^*) = \min\{\bar{e}_1(p_1, u^*), \bar{e}_2(p_2, u^*)\}$$

where $\bar{e}$ is the conditional expenditure function, and $k$ again indexes the chosen (i.e. cost minimising) good, i.e. $k = 1$ if $\bar{e}_1 \leq \bar{e}_2$, or $k = 2$ otherwise. S&R apply the conditional expenditure function to Shephard’s lemma in the following manner, deriving a notion of conditional compensated demand:

$$\frac{\partial e}{\partial p_1} = \begin{cases} \frac{\partial \bar{e}_1}{\partial p_1} = \bar{x}_1^c & \text{if } k = 1 \\ \frac{\partial \bar{e}_2}{\partial p_1} = 0 & \text{if } k = 2 \end{cases}$$

$$\frac{\partial e}{\partial p_2} = \begin{cases} \frac{\partial \bar{e}_1}{\partial p_2} = 0 & \text{if } k = 1 \\ \frac{\partial \bar{e}_2}{\partial p_2} = \bar{x}_2^c & \text{if } k = 2 \end{cases}$$

---

\[^2\text{In an analogous fashion to (3), let us arbitrarily assign } k = 1 \text{ to the case } \bar{e}_1 = \bar{e}_2.\]
where the \( c \) superscript distinguishes the conditional compensated demand \( \bar{x}_i^c \) for good 1 in (4) from the conditional uncompensated demand \( \bar{x}_i \) for good 1 in (3), and similarly for good 2.

### 2.3 Discrete-continuous demand, in practice

Arising from (1), S&R operationalise the demand for good 1 in terms of the following constructs (their equation 3.16):

\[
\begin{align*}
\chi_1(p_1, p_2, y) &= \delta_1(p_1, p_2, y) \bar{x}_1(p_1, p_2, y) \\
\chi_1^c(p_1, p_2, u) &= \delta_1^c(p_1, p_2, u) \bar{x}_1^c(p_1, p_2, u)
\end{align*}
\]

where \( \delta_1 \) and \( \delta_1^c \) are uncompensated and compensated discrete choice indexes respectively, \( \delta_1 = \delta_1^c = 1 \) if \( k = 1 \) (i.e. if \( \bar{v}_1 \geq \bar{v}_2 \), \( \bar{e}_1 \leq \bar{e}_2 \), and good 1 is therefore chosen) or \( \delta_1 = \delta_1^c = 0 \) if \( k = 2 \), and \( \bar{x}_1 \) and \( \bar{x}_1^c \) are (conditional) uncompensated and compensated demands respectively. In other words:

\[
\begin{align*}
\delta_1(p_1, p_2, y) &= \begin{cases} 1 & \bar{v}_1(p_1, y) \geq \bar{v}_2(p_2, y) \\ 0 & \text{otherwise} \end{cases} \\
\delta_1^c(p_1, p_2, u) &= \begin{cases} 1 & \bar{e}_1(p_1, u) \leq \bar{e}_2(p_2, u) \\ 0 & \text{otherwise} \end{cases}
\end{align*}
\]

Although (5), (6) and (7) are couched in terms of good 1, this is simply a presentational convenience; the same analysis could be applied to good 2. We shall adopt this convenience of focussing on good 1 for the remainder of the paper.

Developing the analysis further, S&R translate their model of discrete-continuous demand from the individual level to the aggregate. More specifically, for a
population of $N$ individuals, they define uncompensated and compensated aggregate demands for good 1, respectively:

$$X_1 = \sum_{i=1}^{N} x_1^i = \sum_{i=1}^{N} \delta_1^i (p_1, p_2, y') x_1^i (p_1, p_2, y')$$

(8)

$$X_1^c = \sum_{i=1}^{N} x_1^{c,i} = \sum_{i=1}^{N} \delta_1^{c,i} (p_1, p_2, u') x_1^{c,i} (p_1, p_2, u')$$

(9)

Thus the aggregate uncompensated and compensated demands arise from the summation of the respective individual-level demands over individuals. In aggregating across individuals, the following identities must also hold in relation to discrete choice indexes:

$$\sum_{i=1}^{N} \sum_{j=1}^{2} \delta_1^j (p_1, p_2, y) = \sum_{i=1}^{N} \sum_{j=1}^{2} \delta_1^{c,j} (p_1, p_2, u) = N$$

and budgets:

$$y = \sum_{i=1}^{N} y'$$

(10)

Mindful perhaps of the flexibility that it brings when aggregating surpluses across individuals (see the later discussion in section 6), S&R re-couch the demands (8) and (9) in terms of a ‘representative consumer’, with uncompensated and compensated demands $\bar{x}_1$ and $\bar{x}_1^c$ respectively, and budget $y$ (i.e. omitting the $i$ superscript of the earlier notation). Furthermore, the discrete choice index $\delta_1^i$ is, for the representative consumer, replaced by the discrete choice probability $\pi_1$. Whilst S&R are not explicit about the derivation of the latter, it would seem uncontroversial to suggest that probability arises from the enumeration of the discrete choice index across the population, i.e.

$$\pi_1 = \frac{\sum_{i=1}^{N} \delta_1^i}{N}$$

(11)

where $0 \leq \pi_1 \leq 1$. Using this apparatus of a representative consumer, S&R re-state the aggregate demands (8) and (9), thus:
\[ X_i = \sum_{j=1}^{N} x_i^j (p_1, p_2, y' = y, y' = y') = \sum_{j=1}^{N} \left( \delta_i^j (p_1, p_2, y') \tilde{x}_i^j (p_1, p_2, y') \right) \]

\[ = \sum_{j=1}^{N} \delta_i^j (p_1, p_2, y) \tilde{x}_i^j (p_1, p_2, y) \]

\[ = N \pi_i (p_1, p_2, y) \tilde{x}_i (p_1, p_2, y) \] \hspace{1cm} (12)

\[ X_i^c = \sum_{j=1}^{N} x_i^{c,j} (p_1, p_2, u') = \sum_{j=1}^{N} \left( \delta_i^{c,j} (p_1, p_2, u') \tilde{x}_i^{c,j} (p_1, p_2, u') \right) \]

\[ = N \pi_i^c (p_1, p_2, \{u'\}) \tilde{x}_i^c (p_1, p_2, \{u'\}) \] \hspace{1cm} (13)

and the aggregate budget:

\[ Y = N y \] \hspace{1cm} (14)

Note, with reference to (12) and (8), that (11) holds only if the conditional uncompensated demand for each individual is the same (hence the notion of a representative consumer), that is, if:

\[ \tilde{x}_i^j (p_1, p_2, y') = \tilde{x}_i (p_1, p_2, y) \]

since:

\[ \pi_i (p_1, p_2, y) = \frac{\sum_{j=1}^{N} \delta_i^j (p_1, p_2, y') \tilde{x}_i^j (p_1, p_2, y')}{N \tilde{x}_i (p_1, p_2, y)} \]

This is especially true for the case \( \tilde{x}_1 = \tilde{x}_2 = 1 \). The same rationale applies to the compensated demands (13) and (9).

Having replaced the 0/1 discrete choice index at the individual level with a discrete choice probability at the aggregate level (i.e. applying to the representative consumer), we are called upon to revise our interpretation of \( X_i \) and \( X_i^c \). In particular, it would seem appropriate to interpret \( X_i \) and \( X_i^c \) from (12) and (13) respectively as 'expected' demands. This reflects the fact that, if both
choice probability and conditional demand are non-zero for both goods 1 and 2 (i.e. \( \pi_1, \pi_2 > 0 \) and \( \tilde{x}_1, \tilde{x}_2 > 0 \)), it must be the case that \( x_1 x_2 > 0 \), contrasting with the constraint \( x_1 x_2 > 0 \) applying to (1).

Moreover, this practical implementation of S&R's model of discrete-continuous demand serves to align the utility maximisation problem closer to the conventional presentation in terms of continuous demand. That is to say, for the representative individual, Roy's identity and Shephard's lemma would adopt the following forms:

\[
- \frac{\partial v}{\partial p_1} = \pi_1 (p_1, p_2, y) \tilde{x}_1 (p_1, p_2, y) \\
- \frac{\partial v}{\partial p_2} = \pi_2 (p_1, p_2, y) \tilde{x}_2 (p_1, p_2, y)
\]

\[
\frac{\partial e}{\partial p_1} = \pi_1^c (p_1, p_2, \{u\}) \tilde{x}_1^c (p_1, p_2, \{u\}) \\
\frac{\partial e}{\partial p_2} = \pi_2^c (p_1, p_2, \{u\}) \tilde{x}_2^c (p_1, p_2, \{u\})
\]

(15) (16)

2.4 Eliciting a probabilistic demand model

It is worth remarking that (15) and (16) are routinely exploited in practical demand models, for example in the transportation sector. A popular assumption (e.g. Jara-Díaz & Farah, 1988, eq. 60) however, is that the conditional demands for both goods 1 and 2 are given by a common constant (i.e. \( \tilde{x}_1 = \tilde{x}_2 = \tilde{x} \), noting that \( \tilde{x}_1 = \tilde{x}_1^c \) and \( \tilde{x}_2 = \tilde{x}_2^c \) will hold in equilibrium; we shall formalise the latter proposition in the following section). For example, Hau (1985) observes that: ‘It is in the nature of discrete travel choice models that the total number of trips is assumed fixed’ (p480). Of particular interest to the present paper is the case \( \tilde{x}_1 = \tilde{x}_2 = 1 \), which would apply if the analysis were restricted to a single unit of consumption. As Jara-Díaz & Farah (1988) (their equations 53-57) recognise, in this case, Roy’s identity implies that the conditional marginal utilities of price and income should be equal, and common to both goods 1 and 2, that is:
\[
\frac{\partial \tilde{v}_1}{\partial p_1} = \frac{\partial \tilde{v}_1}{\partial y} = \frac{\partial \tilde{v}_2}{\partial p_2} = \frac{\partial \tilde{v}_2}{\partial y}
\]

Ibáñez & Batley (2010) derive the same result in two alternative ways, using Lagrangian methods and the envelope theorem (the latter being the more general in that it can straightforwardly accommodate multiple constraints). We shall further discuss this equality between the conditional marginal utilities of price and income when discussing implications for model specification in section 4.

Moreover if, having restricted consumption to a single unit, we accept the functional representation of the discrete choice index in terms of prices, income and utility (5, 6), and admit the process of aggregating deterministic individual-level discrete choice indices to yield a probabilistic index for the population (11), then we arrive at the notion of a probabilistic demand function. Indeed, Hau (1985, 1987) adopts this perspective at the outset, deriving the probabilistic demand function from both primal and dual approaches. Since the detailed specification of these probabilistic demands will be an important consideration, the Annex to the present paper outlines Hau’s derivation in full.

3 THE SLUTSKY EQUATION FOR DISCRETE-CONTINUOUS DEMAND

3.1 Introduction

As is widely understood and accepted, the Slutsky equation (Slutsky, 1915) embodies a fundamental relationship between individual-level uncompensated (Marshallian) and compensated (Hicksian) demand functions. Within the Hicksian analysis, a price change influences demand simply through a ‘substitution effect’, whereas within the Marshallian analysis, the change in demand corresponding to a price change can be dissected into both a substitution effect and an ‘income effect’. Following S&R, the present paper will derive the Slutsky equation for the
case of discrete-continuous demand, contrasting with the literature’s usual focus upon the continuous demand.

With reference to the primal problem, S&R specify probabilistic and conditional demands as functional on prices and income in the manner of (12). A minor distinction from S&R (specifically their equation 3.24) is that (12) specifies demand as functional on the income of the representative consumer $y$ rather than aggregate income $Y$ (although, since we assume $Y = Ny$, the dependencies on $Y$ and $y$ are in fact equivalent). Moreover, (12) can be seen as a statement of aggregate expected demand, as the product of probabilistic demand and conditional demand, further multiplied by the number of individuals in the population.

S&R assume that, in equilibrium, uncompensated probabilistic and conditional demands are equivalent to their compensated counterparts, specifically\textsuperscript{3}:

$$
\bar{x}_i(p_1, u) = \bar{x}_i(p_1, \tilde{e}_i(p_1, u))
$$

(17)

$$
\pi_i(p_1, p_2, \{u^i\}) = \pi_i(p_1, p_2, \{\min\{\tilde{e}_1(p_1, u'), \tilde{e}_2(p_2, u')\}\}) = \pi_i(p_1, p_2, \{\varepsilon_i(p_1, p_2, u')\})
$$

(18)

Equation (18) offers a probabilistic analogue to the conventional equilibrium conditions in continuous demand theory. Note that (17) is functional upon the expenditure and utility of the representative consumer, whereas (18) indexes expenditure and utility by individual. The subsequent discussion will proceed by first deriving the Slutsky equation separately for both the conditional demand $\bar{x}_i$

\textsuperscript{3} We should qualify this assertion by noting that S&R, while adopting the notion of equilibrium, in some instances (e.g. their equation 3.24) specify $\bar{x}_i$ to be functional on both $p_1$ and $p_2$. We suggest that this is a typographical error. Intuition tells us that, once good 1 has been chosen over good 2, the demand for good 1 (conditional upon having been chosen) will be functional solely upon own price. Indeed, with reference to S&R’s equation 3.7, which defines $\tilde{e}_i$ as a function of $p_1$ and $u$, it is clear that $\bar{x}_i$ cannot be functional upon $p_2$. 

13
and the probabilistic demand $\pi_1$, before then deriving the Slutsky equation for the aggregate expected demand $X_1$, which arises from the product of $\hat{x}_i$ and $\pi_1$.

### 3.2 Slutsky equation for the conditional demand

From the equilibrium condition for the conditional demand (17), we can derive the Slutsky equation:

$$\frac{\partial \hat{x}_c}{\partial p_i} = \frac{\partial \hat{x}_1}{\partial p_i} + \frac{\partial \hat{x}_1}{\partial \pi_1} \frac{\partial \pi_1}{\partial p_i}$$

(19)

Applying Shephard’s lemma (4) for the conditional demand, and assuming that $\hat{x}_c = \hat{x}_1$ in equilibrium, we can simplify (19) thus:

$$\frac{\partial \hat{x}_c}{\partial p_i} = \frac{\partial \hat{x}_1}{\partial p_i} + \frac{\partial \hat{x}_1}{\partial \pi_1} \hat{x}_1$$

(20)

which is consistent with S&R’s equation (3.18). As is widely understood, the first element on the right hand side of (20) embodies the ‘substitution effect’ of a change in $p_i$, whilst the second element (i.e. product of two terms) embodies the ‘income effect’.

### 3.3 Slutsky equation for the probabilistic demand

Proceeding to the equilibrium condition for the probabilistic demand (18), we can derive the analogous Slutsky equation. Thus, taking derivatives on both sides of the identity, applying the chain rule for derivatives, and acknowledging that conditional expenditure is independent of the price of all alternatives apart from the one that the demand refers to, we have that:
\[
\frac{\partial \pi_i^c(p_1, p_2, \{u^i, \ldots, u^N\})}{\partial p_1} = \frac{\partial \pi_i(p_1, p_2, \{y^i\})}{\partial p_1} \\
+ \sum_{i=1}^{N} \left( \frac{\partial \pi_i(p_1, p_2, \{\hat{e}_{ki}^i(p_k, u^i), \ldots, \hat{e}_{ki}^i(p_k, u^i)\})}{\partial \hat{e}_{ki}^i(p_k, u^i)} \cdot \frac{\partial \hat{e}_{ki}^i(p_k, u^i)}{\partial p_1} \right) \\
+ \sum_{i=1}^{N} \left( \frac{\partial \pi_i(p_1, p_2, \{y^i\})}{\partial y^i} \cdot \delta_{\hat{e}_{ki}^i(p_k, u^i)}(p_1, p_2, u^i) \cdot \xi_{\hat{e}_{ki}^i(p_k, u^i)}(p_1, p_2, u^i) \right)
\]

(21)

Digressing slightly, note that the assumption of identical incomes associated with the representative consumer ensures that the following result holds in equilibrium:

\[
\sum_{i=1}^{N} \left( \frac{\partial \pi_i(p_1, p_2, \{y^i\})}{\partial y^i} \right)_{y^i=\hat{e}_{ki}^i(p_k, u^i)} = \frac{\partial \pi_i(p_1, p_2, y)}{\partial y} \bigg|_{y=\hat{e}_{ki}^i(p_k, u^i)}
\]

(22)

where \( y = \{\hat{e}_{ki}^i(p_k, u^i)\} \) is meaningful given that the \( N \) elements in \( \{\hat{e}_{ki}^i(p_k, u^i)\} \) have identical values at equilibrium. By contrast, there is no need to impose identical utility levels for all individuals.

Making use of (22) and the definition (13) we have that:
\[
\frac{\partial \pi^c_i (p_1, p_2, \{u', \ldots, u^N\})}{\partial p_i} = \frac{\partial \pi_i (p_1, p_2, \{y\})}{\partial p_i} \bigg|_{y' = \delta_i (p, u, u')} + \frac{\partial \pi_i (p_1, p_2, y)}{\partial y} \bigg|_{y = \delta_i (p, u, u')} \cdot X_i (p_1, p_2, \{u'\})
\]

Equation (23) closely resembles S&R’s equation 3.25. We can make further progress towards aligning our analysis with S&R if we suppress the functional dependences, acknowledge that in equilibrium it must hold that \( X_i^c = X_i \) (this follows from (17) and (18)), and admit the following equality between derivatives taken with respect to individual income and aggregate income:

\[
\frac{\partial \pi_1 (p_1, p_2, y)}{\partial y} = \frac{\partial \pi_1 (p_1, p_2, Y)}{\partial Y}
\]

On this basis, we arrive finally at S&R’s Slutsky equation for the probabilistic demand:

\[
\frac{\partial \pi^c_i}{\partial p_i} = \frac{\partial \pi_1}{\partial p_i} + \frac{\partial \pi_1}{\partial Y} X_i
\]

(24)

and in so doing expose the assumptions inherent within S&R’s derivation. In an analogous manner to (20), we can interpret the two elements on the right-hand side of (24) as the ‘substitution effect’ and ‘income effect’ respectively.

### 3.4 Slutsky equation for the aggregate demand

Drawing together the conditional and probabilistic demands, we are now equipped to derive the Slutsky equation for the aggregate (expected) demand, as follows:
\[
\frac{\partial X_i^c}{\partial p_i} = N \left( \frac{\partial \pi_i^c}{\partial p_i} \tilde{X}_i^c + \pi_i^c \frac{\partial \tilde{X}_i^c}{\partial p_i} \right)
\]
\[
= N \tilde{X}_i^c \frac{\partial \pi_i^c}{\partial p_i} + N \pi_i^c \left( \frac{\partial \tilde{X}_i^c}{\partial p_i} + \tilde{X}_i \frac{\partial \tilde{X}_i}{\partial y} \right)
\]
\[
= \frac{X_i^c}{\pi_i^c} \frac{\partial \pi_i^c}{\partial p_i} + N \pi_i^c \left( \frac{\partial \tilde{X}_i^c}{\partial p_i} \left( \frac{X_i}{N \pi_1} \right) + N \tilde{X}_i \frac{\partial \tilde{X}_i}{N \pi_1} \left( \frac{X_i}{N \cdot \pi_1} \right) \right)
\]
\[
= \frac{X_i^c}{\pi_i^c} \left( \frac{\partial \pi_i^c}{\partial p_i} \left( \frac{p_1, p_2, \{y'\}}{\{y'\}} \right) \right) + \frac{X_i^c}{\pi_i^c} \frac{\partial \pi_i^c}{\partial (p_1, p_2, Y)} + N \pi_i^c \left( \frac{1}{N \pi_1} \frac{X_i}{\partial p_i} \left( \frac{X_i}{\pi_1} \right) + N \frac{X_i}{N \pi_1} \left( \frac{\partial X_i}{\partial Y} - \frac{X_i}{\pi_1} \frac{\partial \pi_1}{\partial Y} \right) \right)
\]
\[
= \frac{X_i^c}{\pi_i^c} \left( \frac{\partial \pi_i^c}{\partial p_i} \left( \frac{X_i}{\pi_1} \right) + N \pi_i^c \left( \frac{1}{N \pi_1} \frac{X_i}{\partial p_i} + N \frac{X_i}{N \pi_1} \left( \frac{\partial X_i}{\partial Y} - \frac{X_i}{\pi_1} \frac{\partial \pi_1}{\partial Y} \right) \right) \right)
\]
\[
= \frac{\partial X_i}{\partial p_i} + \frac{X_i}{\pi_1} \left( \frac{\partial X_i}{\partial Y} - (1 - \pi_1) \frac{X_i}{\pi_1} \frac{\partial \pi_1}{\partial Y} \right)
\]

(25)

Note that the second equality of (25) employs the Slutsky equation for the conditional demand (20), that the third equality employs the definitions of aggregate demand (12) and income (13), and that the fourth equality makes use of both the Slutsky equation for probabilistic demand (24) and the following identities:

\[
\frac{\partial \tilde{X}_i}{\partial p_i} = \frac{\partial}{\partial p_i} \left( \frac{X_i}{N \pi_1} \right) = \frac{1}{N \pi_1} \left( \frac{\partial X_i}{\partial p_i} - \frac{X_i}{\pi_1} \frac{\partial \pi_1}{\partial p_i} \right)
\]
\[
\frac{\partial \tilde{X}_i}{\partial y} = \frac{\partial}{\partial y} \left( \frac{X_i}{N \pi_1} \right) = \frac{1}{N \pi_1} \left( \frac{\partial X_i}{\partial y} - \frac{X_i}{\pi_1} \frac{\partial \pi_1}{\partial y} \right)
\]
\[
\frac{\partial \tilde{X}_i}{\partial Y} = \frac{\partial}{\partial Y} \left( \frac{X_i}{N \pi_1} \right) = \frac{1}{N \pi_1} \left( \frac{\partial X_i}{\partial Y} - \frac{X_i}{\pi_1} \frac{\partial \pi_1}{\partial Y} \right)
\]
The final equality of (25) changes variables to consider derivatives with respect to total income rather than individual income. Note that (25) replicates S&R's Slutsky equation for the aggregate demand (their equation (3.26)).

We shall close this section by challenging an assertion made by S&R in footnote 26 of their paper. This footnote attempts to distinguish the approach of deriving expected demand (which S&R claim to follow themselves, using the likes of (12) and (13)), from an alternative approach (followed by Hau (1985, 1987) and Domencich & McFadden (1975)) of deriving demand from expected utility (see the Annex). Whilst such a distinction would seem valid if considering the expected demand in full, the same does not apply when restricting the conditional demand to $\tilde{x}_1 = \tilde{x}_2 = 1$. In this case, the expected demand becomes simply the probabilistic demand, and the derivation of the latter entails a need to take expectations of utility (Hau, 1985, 1987). Moreover, the derivation of the probabilistic demand implies specific assumptions on the formulation of deterministic utility; this will be the subject of the following section.

4 IMPLICATIONS OF THEORY FOR THE EMPIRICAL SPECIFICATION OF DISCRETE CHOICE MODELS

4.1 Introduction

As sections 2 and 3 have acknowledged, if we restrict S&R to the case where $\tilde{x}_1 = \tilde{x}_2 = 1$, then we elicit the notion of a probabilistic demand function. An interest which is developed by Hau (1985, 1987), but not by S&R, is the theoretical derivation of this function from the dual problem of utility maximisation and cost minimisation. The present section will show that the derivation of the probabilistic demand function implies particular requirements on model specification.
4.2 Econometric specification of the probabilistic demand

In order to support the subsequent discussion, it will be helpful to assign further form to the specification of probabilistic demand, in the following respects.

First, and following convention in discrete choice modelling, let us specify the conditional indirect utility for good 1:

\[ \tilde{v}_i(p, y) = W_i(p, y) + \varepsilon_i^i \quad i = 1, ..., N \]  

(26)

where \( W_i \) is the deterministic utility of good 1, the form of which is common to all individuals, and \( \varepsilon_i^i \) is a random error term, which is specific to good 1 and individual \( i \). The above specification is a slightly simplified form of that assumed by S&R (their equation 5.1). Note that deterministic utility is assumed to be functional upon price as well as income of the representative individual, whereas random error is independent of price and income.

Second, let us relate utility to the uncompensated probabilistic demand through the apparatus of the Random Utility Model (RUM) (Marschak, 1960; Block & Marschak, 1960) thus:

\[ \pi_1 = \Pr \{ \tilde{v}_1 \geq \tilde{v}_2 \} = \Pr \{ W_1 + \varepsilon_1 \geq W_2 + \varepsilon_2 \} = \Pr \{ W_1 - W_2 \geq \varepsilon_1 - \varepsilon_2 \} = \varphi(W_1 - W_2) \]  

(27)

where \( \varphi \) is the distribution function of \( \varepsilon_1 - \varepsilon_2 \). RUM itself embodies several properties (see for example Daly & Zachary (1978) for a discussion), but one which will be especially relevant to the subsequent discussion is ‘translational invariance’, meaning that probability is robust to any (common) increasing linear transformation of \( W_1 \) and \( W_2 \). It is the difference between the deterministic utilities of goods 1 and 2 \((W_1 - W_2)\) that influences probability, not the absolute utilities.

Having introduced (26) and (27), it should be clarified that the present paper will, in exploring the implications of S&R for the practical specification of probabilistic
demand models, focus upon deterministic utility $W_i(.)$ rather than random error $\epsilon_i$. That is not to overlook important specification issues concerning random error, which have been considered by a number of previous authors; for a recent review and reassessment of those issues, see Ibáñez (2007).

4.3 Four model specification assumptions

Following the usual conventions of dual consumption problems, Hau (1985, 1987) derives the probabilistic demand function via both Roy’s identity and Shephard’s lemma. This derivation, which is reproduced in the Annex, imposes four assumptions on the specification of deterministic utility; these assumptions will be the subject of the subsequent discussion. The derivation also makes a preliminary assumption concerning the random error term in (26), specifying this to be IID Gumbel; this assumption will not substantively affect the assumptions concerning deterministic utility, to which we now turn.

First and foremost, in the restricted case of the probabilistic demand, the outcome of Roy’s identity for the conditional demand must be $\bar{x}_i = \bar{x}_2 = 1$, which implies:

- **Assumption I**: for each alternative, equivalence (in absolute terms) between the conditional marginal utilities of income and price, i.e. $-\partial W_1 / \partial p_1 = \partial W_1 / \partial y, -\partial W_2 / \partial p_2 = \partial W_2 / \partial y$

Now focussing more specifically upon Hau’s derivation, intuition tells us that the outcome of Roy’s identity for the probabilistic demand should be $\pi_1$. As the Annex demonstrates, this result implies that Assumption I should be combined with:

- **Assumption II**: common conditional marginal utility of income across alternatives, i.e. $\partial W_1 / \partial y = \partial W_2 / \partial y = \lambda$
If, for empirical purposes, the probabilistic demand were specified in the manner of McFadden’s (1981) ‘additive income RUM’ (AIRUM) then, given the property of translational invariance described in section 4.2, Assumption II would imply that demand is invariant to changes in income. In this case, the ‘income effect’ identified in (24) would be zero. Note furthermore that Assumptions I and II, taken together, imply a further assumption:

- **Assumption III**: common conditional marginal utility of price across alternatives, i.e. $\frac{\partial W_1}{\partial p_1} = \frac{\partial W_2}{\partial p_2}$

Whilst the norm in transport economic practice is to assume $\frac{\partial W_1}{\partial p_1} = \frac{\partial W_2}{\partial p_2}$, implying a generic parameter for the conditional marginal utility of price, a number of papers could be cited (e.g. Swait & Ben-Akiva, 1987; de Jong et al., 2003; and Hess et al., 2007) that appear to offer empirical support for $\frac{\partial W_1}{\partial p_1} \neq \frac{\partial W_2}{\partial p_2}$. Section 6.4 of the present paper will argue that the latter inequality implies particular theoretical properties on the part of the probabilistic demand. Note that Assumptions I and III (together) similarly imply II, but that II and III do not imply Assumption I.

Finally, Hau’s derivation of the probabilistic demand imposes a fourth assumption:

- **Assumption IV**: the conditional marginal utility of income, $\lambda$, is independent of the prices of goods 1 and 2, i.e. $\frac{\partial \lambda}{\partial p_1} = \frac{\partial \lambda}{\partial p_2} = 0$

Whereas Assumption III is sufficient to eliminate income effects in the AIRUM model, Assumption IV eliminates income effects across the class of RUM models more generally. We shall consider this property in greater detail in the course of our discussion of welfare in section 6.
5 THEORETICAL PROPERTIES OF THE PROBABILISTIC DEMAND

5.1 Introduction

As was noted in section 1, and is widely understood and accepted (see for example Deaton & Muellbauer (1980) for a general discussion), demand functions should observe four properties, namely ‘adding up’, ‘negativity’, ‘homogeneity’ and ‘symmetry’. The purpose of the present section is to consider the extent to which the probabilistic demand arising from S&R complies with the aforementioned properties. In the course of this discussion we shall draw upon the four assumptions of the previous section which, as we have already acknowledged, are inherent within the derivation of the probabilistic demand function. Intuition suggests that there should be correspondence between the assumptions imposed on model specification, and the theoretical properties of the resulting demand functions; in what follows, we shall confirm whether or not this is indeed the case.

In general, the four theoretical properties are relevant to individual-level demands, and may not readily translate to aggregate demand (a notable example being symmetry). However, having adopted the notion of a representative consumer, the distinction between individual and aggregate level demands becomes, in the case of S&R at least, a rather moot point. On this basis, let us consider the theoretical properties of the aggregate demand functions (12) and (13), albeit for the case where \( N=1 \) (i.e. notionally applying to a single representative individual).

5.2 Adding up

The ‘adding up’ property requires that the total value of both the Marshallian and Hicksian demands is total expenditure. Formally:
$p_1x_1(p_1, p_2, y) + p_2x_2(p_1, p_2, y) = p_1x_1^c(p_1, p_2, u) + p_2x_2^c(p_2, p_2, u) = y$

Since S&R specify the budget constraint in (1) as an equality, it quickly becomes clear that all budget is exhausted and adding up is imposed by definition. We will not therefore devote this property further attention.

5.3 Negativity

As is widely understood and accepted, a fundamental requirement on demand functions is compliance with the so-called ‘integrability’ conditions (Hurwicz & Uzawa, 1971; Houthakker, 1950; Samuelson, 1950). These conditions require the compensated demand function to exhibit a symmetric negative semi-definite substitution matrix.

We shall deal with symmetry in section 5.5. The present section considers the ‘negativity’ property, which requires that the partial derivatives of the compensated demand with respect to own price are less than or equal to zero, i.e.

$$\frac{\partial x_i^c}{\partial p_i} \leq 0$$  \hspace{1cm} (28)

Although (28) is stated in terms of good 1, the same property should apply analogously to good 2. Moreover, an implication of negativity is that, for both goods 1 and 2, any (positive) income effect on demand should be outweighed by a (negative) substitution effect. Substituting for the Slutsky equation from the conditional (20) and probabilistic (24) demands, we can re-state negativity as follows:

$$\frac{\partial x_i^c}{\partial p_i} = \left[ \left( \frac{\partial \pi_i}{\partial p_i} + \frac{\partial \pi_i}{\partial y} \right) \bar{x}_i + \pi_i^c \left( \frac{\partial \bar{x}_i}{\partial p_i} + \frac{\partial \bar{x}_i}{\partial y} \right) \right] \leq 0$$

Then admitting the econometric specification (26) and (27), we can expand $\partial \pi_i/\partial p_i$ and $\partial \pi_i/\partial y$ thus:
\[
\frac{\partial x_i^c}{\partial p_i} = \left[ \frac{\partial \pi_i}{\partial p_i} \frac{\partial W_i}{\partial p_i} + \frac{\partial \pi_i}{\partial W_i} \frac{\partial W_i}{\partial x} x_i \right] \chi_i^c + \pi_i \left[ \frac{\partial \chi_i}{\partial p_i} + \frac{\partial \chi_i}{\partial y} \right] \leq 0
\] 

(29)

Given the generality of (29), it may be difficult on initial inspection to discern insights for the practical specification of demand models. However, we can make progress if we consider restricted cases of (29). More specifically, let us follow a similar tack to section 2.4, by assuming that the conditional demand for good 1 is unitary (i.e. \(\chi_i = \chi_i^c = 1\)) and independent of price and income (i.e. \(\partial \chi_i / \partial p_i, \partial \chi_i / \partial y = 0\)). In this case the aggregate (expected) demand effectively becomes the probabilistic demand, and (29) simplifies to:

\[
\frac{\partial x_i^c}{\partial p_i} = \left[ \frac{\partial \pi_i}{\partial W_i} \frac{\partial W_i}{\partial p_i} + \frac{\partial \pi_i}{\partial W_i} \frac{\partial W_i}{\partial y} \pi_1 \right] \leq 0
\] 

(30)

Further simplification of (30) yields a final statement of the requirements for negativity, thus:

\[
-\frac{\partial W_i}{\partial p_i} \geq \frac{\partial W_i}{\partial y} \pi_1
\] 

(31)

Since the restricted case \(\chi_i = \chi_i^c = 1\) implies that Assumption I must hold, and \(0 \leq \pi_1 \leq 1\) by definition, we conclude that negativity is assured.

As a closing remark, it is appropriate to emphasise that the negativity condition (31) refers to a restricted case of the aggregate demands (12) and (13), albeit one that is commonly applied in practical demand studies. Unfortunately, the negativity condition (29) applying to the more general case lends itself less readily to the elicitation of clear prescription for model specification. The same caveat will apply to the discussions of homogeneity and symmetry that follow.
5.4 Homogeneity

As is well established in the literature on ‘homogeneity’, the compensated demands should be homogenous of degree zero in prices, and the uncompensated demands should be homogenous of degree zero in prices and income. With reference to Euler’s Theorem, the implication follows that the marginal utility of income should be homogenous of degree minus one in prices. These properties arise from the linear budget constraint in (1), and ensure the absence of ‘money illusion’. More formally, for a positive factor \( \theta > 0 \), it must in equilibrium hold that:

\[
x_i^c(\theta p_1, \theta p_2, u) = x_i^c(p_1, p_2, u) = x_i(\theta p_1, \theta p_2, \theta y) = x_i(p_1, p_2, y)
\]

(32)

That is to say, if all prices and income are increased by a common factor then the demand for good 1 is unchanged. The same property should also apply to good 2.

An alternative formalisation of homogeneity is in terms of the following identity (where for brevity we restrict attention to the uncompensated demand, although a similar identity could be formulated in respect of the compensated demand):

\[
\frac{\partial x_i}{\partial p_1} p_1 + \frac{\partial x_i}{\partial p_2} p_2 + \frac{\partial x_i}{\partial y} y = 0
\]

(33)

For a derivation of (33) from (32), see for example DeSerpa (1971), although note the typographical error in his equation (2.16), which is corrected above. With reference to our aggregate demands (12) and (13), and also drawing upon (26) and (27), let us derive the various components of (33), thus:

\[
\begin{align*}
\frac{\partial x_i}{\partial p_1} &= \frac{\partial \pi_1}{\partial W_1} \frac{\partial W_1}{\partial p_1} \hat{x}_i + \pi_1 \frac{\partial \hat{x}_i}{\partial p_1} \\
\frac{\partial x_i}{\partial p_2} &= \frac{\partial \pi_1}{\partial W_2} \frac{\partial W_2}{\partial p_2} \hat{x}_i + \pi_1 \frac{\partial \hat{x}_i}{\partial p_2} \\
\frac{\partial x_i}{\partial y} &= \left[ \frac{\partial \pi_1}{\partial W_1} \frac{\partial W_1}{\partial y} + \frac{\partial \pi_1}{\partial W_2} \frac{\partial W_2}{\partial y} \right] \hat{x}_i + \pi_1 \frac{\partial \hat{x}_i}{\partial y}
\end{align*}
\]
Substituting back into (33), we have:

\[
\begin{align*}
\frac{\partial \pi_1}{\partial W_1} & + \frac{\partial \pi_1}{\partial \hat{\pi}_1} \frac{\partial}{\partial \hat{\pi}_1} \hat{\pi}_1 + \frac{\partial \pi_2}{\partial W_2} & + \frac{\partial \pi_2}{\partial \hat{\pi}_2} \frac{\partial}{\partial \hat{\pi}_2} \hat{\pi}_2 + \frac{\partial \pi_2}{\partial \hat{\pi}} \frac{\partial}{\partial \hat{\pi}} \hat{\pi} + \frac{\partial \pi_2}{\partial \hat{\pi}_1} \frac{\partial}{\partial \hat{\pi}_1} \hat{\pi}_1 + \frac{\partial \pi_2}{\partial \hat{\pi}_2} \frac{\partial}{\partial \hat{\pi}_2} \hat{\pi}_2 + \frac{\partial \pi_2}{\partial \hat{\pi}} \frac{\partial}{\partial \hat{\pi}} \hat{\pi} + \frac{\partial \pi_2}{\partial \hat{\pi}_1} \frac{\partial}{\partial \hat{\pi}_1} \hat{\pi}_1 + \frac{\partial \pi_2}{\partial \hat{\pi}_2} \frac{\partial}{\partial \hat{\pi}_2} \hat{\pi}_2 + \frac{\partial \pi_2}{\partial \hat{\pi}} \frac{\partial}{\partial \hat{\pi}} \hat{\pi} + \frac{\partial \pi_2}{\partial \hat{\pi}_1} \frac{\partial}{\partial \hat{\pi}_1} \hat{\pi}_1 & \quad \text{for } p_1 + \\
\left( \frac{\partial \pi_1}{\partial W_1} \frac{\partial}{\partial y} \right) & + \left( \frac{\partial \pi_2}{\partial W_2} \frac{\partial}{\partial y} \right) \frac{\partial}{\partial y} y = 0
\end{align*}
\]

Following the same rationale as our discussion of negativity (section 5.3), let us again restrict attention to the probabilistic demand, which enables simplification of (34) thus:

\[
\begin{align*}
\frac{\partial \pi_1}{\partial W_1} p_1 & + \frac{\partial \pi_1}{\partial W_2} p_2 & \quad \text{for } p_1 + \\
\left[ \frac{\partial \pi_1}{\partial W_1} + \frac{\partial \pi_1}{\partial W_2} \frac{\partial}{\partial y} \right] & + \left( \frac{\partial \pi_2}{\partial W_1} + \frac{\partial \pi_2}{\partial W_2} \frac{\partial}{\partial y} \right) y = 0
\end{align*}
\]

Unfortunately, (35) does not immediately yield prescription for the specification of discrete choice models. However, we can make further progress if we adopt Assumptions I and III, given that they are imposed by Hau (1985, 1987) in the course of deriving the probabilistic demand. On this basis, and noting that homogeneity should apply to both goods, (35) can be combined with the analogous conditions for good 2, to yield the following identity:

\[
\begin{align*}
\left( \frac{\partial \pi_1}{\partial W_1} - \frac{\partial \pi_2}{\partial W_2} \right) p_1 & + \left( \frac{\partial \pi_1}{\partial W_2} - \frac{\partial \pi_2}{\partial W_2} \right) p_2 & \quad \text{for } p_1 + \\
\left[ \frac{\partial \pi_1}{\partial W_1} - \frac{\partial \pi_2}{\partial W_2} + \left( \frac{\partial \pi_1}{\partial W_1} - \frac{\partial \pi_2}{\partial W_2} \right) + \left( \frac{\partial \pi_1}{\partial W_2} - \frac{\partial \pi_2}{\partial W_2} \right) \right] y = 0
\end{align*}
\]

It therefore becomes apparent that a key driver in establishing compliance with homogeneity is the matrix of derivatives of probability with respect to deterministic utility, as given by:

\[
\begin{align*}
\frac{\partial \pi}{\partial W} = \begin{bmatrix}
\frac{\partial \pi_1}{\partial W_1} \quad -\frac{\partial \pi_1}{\partial W_2} \\
\frac{\partial \pi_2}{\partial W_1} \quad -\frac{\partial \pi_2}{\partial W_2}
\end{bmatrix}
\end{align*}
\]

As is widely understood (e.g. Steenburgh, 2008), a property of RUM is that each of the rows and columns of the matrix (37) sums to zero. Therefore, if the
probabilistic demand model is specified as RUM then (36) must hold, and homogeneity will be satisfied. Note that this result runs counter to Hau’s (1985) assertion (which he does not formally justify) that AIRUM does not comply with homogeneity.

5.5 Symmetry

Further to the discussion of negativity in section 5.3, a second feature of the integrability conditions is the ‘Slutsky symmetry’ property (Slutsky, 1915). Before proceeding, it should be qualified that symmetry is, strictly speaking, relevant to individual-level demand. Although there is no analogous requirement on the aggregate demand, S&R’s adoption of the representative consumer implies that their demand function will embody consistent properties, whether viewed in aggregate or disaggregate.

Arising from Shephard’s lemma and Young’s theorem, the Slutsky symmetry property imposes the following conditions on the individual-level unconditional compensated demands:

\[
\frac{\partial x_1^c}{\partial p_2} = \frac{\partial}{\partial p_2} \left( \frac{\partial e}{\partial p_1} \right) = \frac{\partial}{\partial p_1} \left( \frac{\partial e}{\partial p_2} \right) = \frac{\partial x_2^c}{\partial p_1}
\]

Substituting for the unconditional compensated demand, using our notion of expected demand (13), and again assuming \( N = 1 \), this property can be restated:

\[
\frac{\partial \pi_1^c \bar{x}_1^c}{\partial p_2} = \frac{\partial \pi_2^c \bar{x}_2^c}{\partial p_1}
\]

Further substituting for the Slutsky equations from the conditional (20) and probabilistic (24) demands, and drawing upon the econometric specification (26) and (27), we can expand the terms of (39) as follows:
In this way, we derive a statement of Slutsky symmetry, which must hold if the uncompensated demands for goods 1 and 2 are to satisfy integrability.

Now adopting the same focus as the earlier discussion of homogeneity, let us restrict (40) to the probabilistic demand, by noting that \( \pi_1 = \pi_1^c \) and \( \pi_2 = \pi_2^c \) hold in equilibrium, and by assuming \( \partial x_1 / \partial p_2^c, \partial x_2 / \partial p_2^c, \partial x_1 / \partial y, \partial x_2 / \partial y = 0 \) and \( \tilde{x}_1 = \tilde{x}_2 = \tilde{x}_2^c = 1 \). If we again impose Assumption I, which would seem to be unavoidable in deriving the probabilistic demand, then symmetry amounts to the following condition:

\[
[1 + \pi_2] \left[ \frac{\partial \pi_1}{\partial W_2} \frac{\partial W_2}{\partial p_2} \right] = [1 + \pi_1] \left[ \frac{\partial \pi_2}{\partial W_1} \frac{\partial W_1}{\partial p_1} \right]
\]  

(41)

In the extant literature, various authors (including S&R themselves and Daly & Zachary, 1978) represent symmetry simply in terms of the cross-partial derivatives of probability with respect to deterministic utility (i.e. \( \partial \pi_1 / \partial W_2 = \partial \pi_2 / \partial W_1 \), referred to as ‘Condition 5’ in Daly & Zachary). Although necessary to ensure that choice probabilities can be generated by a correctly specified multivariate distribution for the random error terms, (41) shows that Condition 5 is not in itself sufficient to ensure Slutsky symmetry.

Indeed, equation (41) reveals that Slutsky symmetry is, in principle, dependent not only on the cross-partial derivatives of probability with respect to deterministic utility, but also the conditional marginal utilities of price for goods 1 and 2, and the market shares. Things become clearer if we also impose Assumption III (remembering that Assumptions I, II and III are all inherent within Hau’s (1985, 1987) derivation of the probabilistic demand), as well as Daly & Zachary’s (1978)
‘Condition 5’ (given the justification outlined above). On this basis, the result emerges that symmetry holds only if \( \pi_1 = \pi_2 = 0.5 \).

Since equi-probability will not be the norm (indeed this could imply a model which has no explanatory power), one might be drawn to conclude that many practical models will fail to comply with Slutsky symmetry. However, there is a final piece of the jigsaw to be added, in the form of Assumption IV. As was acknowledged in section 4.3, Assumptions I, II and III would, in the case of AIRUM, be sufficient to preclude an income effect. The addition of Assumption IV would be sufficient preclude an income effect in any RUM.

Moreover, if we accept Assumptions I-IV, then the symmetry conditions simplify to:

\[
\frac{\partial \pi_1}{\partial W_2} = \frac{\partial \pi_2}{\partial W_1}
\]

and we return thus to the interpretation of symmetry promoted by S&R and Daly & Zachary, which is satisfied by any RUM. We conclude that a (Marshallian) probabilistic demand model specified as RUM will, aside from the artefactual case of \( \pi_1 = \pi_2 = 0.5 \), comply with Slutsky symmetry only in the absence of income effects. In this way, we arrive at a probabilistic version of the conventional result that a Marshallian demand will, in the absence of income effects, comply with the conditions for path independence. We shall develop this point further in the following section.
6 WELFARE MEASURES FROM DISCRETE-CONTINUOUS DEMANDS

6.1 Introduction

S&R’s paper culminates in a derivation of consumer surplus from their model of discrete-continuous demand, but arriving at a result defined entirely in terms of the probabilistic demand. The following discussion will review this derivation, with the aim of exposing S&R’s underlying rationale, and reconciling that rationale with the conclusions on empirical specification and theoretical properties from sections 4 and 5 of the present paper.

6.2 Definitions of Hicksian consumer surplus

S&R begin their discussion of welfare at the individual level, deriving the compensating variation of a change in the price of good 1 from $p^0_1$ to $p^1_1$ as the integral of the unconditional compensated demand:

$$\Delta e = e(p^1_1, p^1_2, u^{0_1}) - e(p^0_1, p^0_2, u^{0_1}) = \int_{p^0_1}^{p^1_1} x^{c_1}_i(p_1, p_2, u^{0_1}) dp_1$$

where the superscript 0 refers to the base case, and $f$ to the forecast case.

Taking the population of consumers as a whole, S&R acknowledge that, whilst facing common prices, different consumers may make different choices and derive different utilities. On this basis, S&R propose an aggregate analogue to the compensating variation, thus:

$$\Delta E = \int_{p^0_1}^{p^1_1} \sum_{i=1}^N \delta^{c_1}_i(p_1, p_2, u^{0_1}) \tilde{x}^{c_1}_i(p_1, p_2, u^{0_1}) dp_1$$

Or replacing the deterministic discrete choice indicator with the probabilistic:
\[ \Delta E = \int_{\mathbb{R}_+} \sum_{i=1}^{N} \pi_{1i}^c (p_1, p_2, u^o_1) \hat{x}_{1i}^c (p_1, p_2, u^o_1) dp_i \]  

(43)

Alternatively, we could adopt the perspective of a representative consumer, in which case (43) could be re-stated:

\[ \Delta E = \int_{\mathbb{R}_+} N \pi_{1i}^c (p_1, p_2, u^o) \hat{x}_{1i}^c (p_1, p_2, u^o) dp_i \]

As section 6.5 will discuss, the notion of representative consumption has particular significance for welfare measurement.

6.3 Isolating consumer surplus specific to the probabilistic demand

Although S&R introduce the concept of consumer surplus in terms of the compensated (Hicksian) demand, in practice they adopt the uncompensated (Marshallian) demand as a working approximation; we shall rehearse the rationale behind this approximation in section 6.4. The present section will discuss a more general point concerning S&R’s interpretation of the demand function when deriving consumer surplus. This discussion will be conducted in terms of the Marshallian, but the substantive point is equally relevant to the Hicksian.

In deriving welfare, S&R integrate over the product of the probabilistic and conditional (Marshallian) demands, which would seem to imply that the relevant demand function for welfare measurement is the expected demand. However, they subsequently apply Roy’s identity to the conditional demand (i.e. following (3)), arriving at the following expression for welfare change:

\[ \Delta E \approx \int_{\mathbb{R}_+} \frac{N}{\lambda (y)} \pi_1 (p_1, p_2, y^0) \frac{\partial W_i}{\partial p_i} \frac{\partial W_i}{\partial y} dp_i \]  

(44)
where we replace $e$ with $y$, which must of course hold in equilibrium, and assume (more strongly than S&R) that $\lambda$ is independent of both $p_1$ and $p_2$. We shall comment further on the functionality of $\lambda$ in section 6.4.

It is clear from (44) that the integral over price for $x_i$ is compatible with the application of Roy’s identity to $\tilde{x}_i$ only if $\pi_i$ is a constant, and therefore independent of price. Indeed, on this point, (44) seems to run counter to advice issued by S&R earlier in their paper: ‘Note that only if the choice probability is independent of price $p$, will the correct answer be obtained by calculating areas for conditional demand curves and multiplying them by the probabilities. Thus, for example, Feldstein and Friedman (1977) are correct in using the latter procedure to calculate welfare effects of changes in the price of health insurance, because in their model the probability of illness is assumed to be exogeneous. However, one must not be misled into thinking their procedure is generalizable to endogenous probabilities’ (footnote 18, p115).

Similarly, the integral over price for $x_i$ is compatible with application of Roy’s identity to $\pi_i$ provided $\tilde{x}_i$ is instead a constant. In particular, if we assume (as we did earlier in section 4) that the conditional demand is unitary, then it follows from (44) that Assumption I must hold, i.e. $-\partial W_i/\partial p_i = \partial W_i/\partial y$. This result is corroborated by Hau (1985, 1987), which implicitly assumes $\tilde{x}_i = 1$, and yields the probabilistic demand $\pi_i$ from Roy’s identity, provided $-\partial W_i/\partial p_i = \partial W_i/\partial y$ (see the Annex to the present paper).

Notwithstanding the above critique, let us proceed with S&R’s analysis, by noting their remark that: ‘… $\pi_i$ depends on its arguments only through the functions $W_1$ and $W_2$’ (p124). This observation is used to justify a change of variable $\omega_1 = W_1(p_i, y)$, and a re-working of (44) which arrives at (essentially) the following expression:
\[ \Delta E \approx -\frac{N}{\lambda(y)}\int_{W_1}^{W_1'} \pi_1(W_1, W_2) dW_1 \]  

(45)

where \( W_i^0 = W_i(p_i^0, y^0) \) and \( W_i' = W_i(p_i', y^0) \).

In short, S&R propose a measure of welfare change (45), in money units, which derives from the integration of the probabilistic demand for a change in indirect deterministic utility associated with a price change. This is the basic intuition behind the log sum construct used widely in public policy analysis, which arises from (45) under an assumption that the random error term in (26) is distributed IID Gumbel. Completing our discussion of welfare, the subsequent sections 6.4 and 6.5 will comment upon two features of (45), namely the adoption of Marshallian consumer surplus as an approximation to the Hicksian, and the aggregation of consumer surplus across individuals.

6.4 Marshallian consumer surplus as an approximation to the Hicksian

Having introduced the notion of Hicksian consumer surplus, S&R introduce a series of assumptions on the basis of ‘purely empirical considerations’ (their terminology), two of which are especially pertinent to our own interests.

- **Assumption V:** ‘…the marginal utility of income, \( \lambda \), is approximately independent of the price…of good 1’ (p124).

- **Assumption VI:** ‘…the discrete goods are sufficiently unimportant to the consumer so that income effects…are negligible, i.e. that the compensated demand…is adequately approximated by the Marshallian demand function’ (p124).

These two assumptions are used to justify S&R’s adoption of Marshallian consumer surplus as an approximation to Hicksian (as illustrated in (44) and (45) above), a practice which is *de rigeur* in applied welfare economics (e.g. Morey,
Against this background, let us consider the rationale for each of these assumptions in more detail.

Assumption V would seem to be motivated by a wish to nullify the 'path dependency' property of Marshallian consumer surplus. If that is indeed the intention, then it would be appropriate to strengthen S&R’s assumption such that it applies to the prices of both goods 1 and 2 (i.e. $\frac{\partial \lambda}{\partial p_1} = \frac{\partial \lambda}{\partial p_2} = 0$). Moreover, as has already been exposed by section 4, a comprehensive assumption (i.e. Assumption IV) of path independence is inherent within Hau’s (1985, 1987) theoretical derivation of the probabilistic demand.

Whilst Hau and S&R agree upon the assumption of path independence, there is a question as to whether this assumption is supportable by theory (as implied by Hau) or empirics (as implied by S&R). On the one hand, we note Batley & Ibáñez’s (2011) rationale that, in the case $x_1 = x_2 = 1$, the response to any price change can be explained entirely in terms of a substitution effect on the probabilistic demand. If we accept this rationale, then Assumption IV might be seen as formalising a universal theoretical property of probabilistic demand models, and Assumption V might be considered unnecessary (on the basis that income effects are empirically irrelevant). If, on the other hand, we countenance the possibility (in principle) of income effects of a price change on probabilistic demand, then Assumption IV might be seen as a restriction of the model to the special case of null income effects, and Assumption V might be seen as an empirical assumption that supports the applicability of the model for that special case. Although we might debate the theoretical and/or empirical motivation for Assumptions IV and V, one point that we can be clear upon is that the log sum metric, which arises from (45) if the random error term in (26) is IID Gumbel, embodies path independence.

Assumption IV implies the property $\lambda = \lambda(y)$, meaning that if prices change with income fixed then the compensating variation will be proportional to the
Marshallian consumer surplus (Samuelson, 1942). Furthermore, this property imposes specific form upon the underlying utility function, which must be homothetic (Silderberg, 1972). Having exposed the presence of homotheticity within S&R’s analysis, this implies clear prescription for model specification; if practitioners wish to adhere to S&R, and in particular apply welfare measures in the form of (45), then models should be specified such that probability is invariant to changes in income. Indeed, this prescription corroborates our conclusion on symmetry (section 5.5), which should come as no surprise given the intimate relationship between the properties of homotheticity and symmetry (see for example Chapter 9 of Varian (1992) and Chapter 3 of Johansson (1987)). In short, practitioners should avoid model specifications which, for a given income change, allow individuals to switch their discrete choice.

Assumption IV is not however sufficient - in continuous demand theory, at least - to justify the adoption of the Marshallian consumer surplus as an approximation to the Hicksian. This is because income could continue to influence the Marshallian demand for goods 1 and 2, albeit demand in fixed proportion. Recognition of this possibility perhaps motivates Assumption VI, which eliminates the influence of income on the Marshallian demand altogether. However, in the case where \( \tilde{x}_1 = \tilde{x}_2 = 1 \), S&R’s model of discrete-continuous demand reduces to our present interest in the probabilistic demand, and it must hold that \( \pi_1 + \pi_2 = 1 \). This property implies that the uncompensated probabilistic demands for goods 1 and 2 cannot be increased (or reduced) in fixed proportion as income changes. That is to say, by restricting our model to the case \( \tilde{x}_1 = \tilde{x}_2 = 1 \), Assumption VI is imposed by construction, and we need make no further provision on the basis of ‘empirical considerations’.
6.5 From the individual to the aggregate

In closing our discussion of welfare, it is appropriate to acknowledge a final complication which arises when aggregating uncompensated surpluses across individuals. As is widely acknowledged in the literature (e.g. Johansson, 1987), if variations in the marginal utility of income across individuals are not accounted for, then the contentious phenomenon will arise that richer individuals are assigned larger welfare weights than poorer individuals. Such considerations perhaps motivate S&R’s adoption of the ‘representative consumer’ (Gorman, 1953; Diewert, 1980), since this is a popular device for avoiding problems of aggregation. Whilst the vast majority of discrete choice modelling applications have been faithful to the representative consumer when implementing (45), some researchers have explicitly accommodated variation in the marginal utility of income by specifying several income segments, each with its own representative consumer (see de Jong et al. (2007) for a review of applications). It is interesting to note that such a specification is not inconsistent with Assumptions I-IV and, by implication, the application of S&R’s welfare measure. That is to say, provided the conditional marginal utilities of both price and income are indexed by income segment, Assumptions I-IV would allow variation in the marginal utility of income across income segments, but impose a null income effect within each segment.

7 SUMMARY AND CONCLUSIONS

The practical context for S&R’s analysis is a demand problem where commodities are available in continuous quantities but only a limited number of varieties. For a simple case involving goods 1 and 2 only, S&R distinguish between two facets of this demand problem, namely discrete choice between the two goods, and the consumption of a continuous quantity conditional upon choice. S&R represent this
demand problem in terms of an uncompensated expected demand function for a representative individual. In implementing this expected demand, there is an important distinction between what S&R term ‘endogenous’ and ‘exogenous’ choice probability. S&R assume the former, which gives rise to distinct demand models for both the conditional and probabilistic components.

For purposes of econometric implementation, S&R specify the probabilistic demand model as RUM, such that probability is related to the conditional utility of goods 1 and 2, and conditional utility is further dissected into deterministic and random components. The primary analytical focus of S&R’s paper is to derive a measure of consumer surplus specific to the probabilistic demand. To this end, an initial need is to elicit the probabilistic demand from the expected demand. As the present paper has shown, the procedure of eliciting the probabilistic demand imposes a particular assumption on the specification of deterministic utility. Taking good 1 as an example, it is necessary to assume \( -\partial W_1 / \partial p_1 = \partial W_1 / \partial y \), meaning that the conditional marginal utilities of price and income must be equal. In this case, the conditional demands would, via Roy’s identity, be restricted to \( \hat{x}_1 = \hat{x}_2 = 1 \), and the expected demand would reduce to the probabilistic demand.

This result confirms the finding of Jara-Díaz & Farah (1988), and is corroborated by related work by Hau (1985, 1987).

Having dissected the probabilistic demand from the conditional demand, an interest in consumer surplus focuses attention upon the workings of the Slutsky equation. The present paper replicated S&R’s derivation of the Slutsky equation for both the conditional and probabilistic elements of demand, exposing the assumptions underpinning this derivation. A feature of S&R’s Slutsky equation for the probabilistic demand is that, in general, the unconditional demand is the outcome of Shephard’s lemma. However, in the restricted case \( \hat{x}_1 = \hat{x}_2 = 1 \), Shephard’s lemma yields instead the probabilistic demand, and S&R’s analysis then coincides with Hau’s (1985, 1987).
For this restricted case, there is an important question as to what assumptions are inherent in the derivation of the probabilistic demand from Shephard’s lemma. According to Hau, the derivation of the probabilistic demand requires the assumptions $-\partial W_i/\partial p_1 = \partial W_i/\partial y$, $\partial W_i/\partial y = \partial W_2/\partial y = \lambda$ (which also then implies $\partial W_i/\partial p_1 = \partial W_2/\partial p_2$) and $\partial \lambda/\partial p_1 = \partial \lambda/\partial p_2 = 0$.

Developing the discussion further, it is interesting to consider the coherence between the assumptions imposed by Hau, and four theoretical properties of the probabilistic demand, namely ‘adding up’, ‘negativity’, ‘homogeneity’ and ‘symmetry’. Given the manner in which S&R specify the budget condition, ‘adding up’ holds by definition. If we accept the assumption $-\partial W_i/\partial p_1 = \partial W_i/\partial y$ (as is inherent in deriving the probabilistic demand), then ‘negativity’ is similarly assured. If we further assume $\partial W_i/\partial p_1 = \partial W_2/\partial p_2$ (also then implying $\partial W_i/\partial y = \partial W_2/\partial y = \lambda$) then ‘homogeneity’ holds. The assumptions applied thus far do not however guarantee ‘symmetry’. This requires the addition of a final assumption $\partial \lambda/\partial p_1 = \partial \lambda/\partial p_2 = 0$, which imposes path independence, eliminates any income effect, and establishes equivalence between the surpluses arising from Marshallian and Hicksian probabilistic demands.

Drawing the preceding discussion together, some clear prescriptions emerge, so as to ensure that S&R’s objective of conducting welfare analysis using only the probabilistic component of the expected demand can be realised in a theoretically-defensible manner.

First and foremost, it should be acknowledged that the welfare measure (45) promoted by S&R - an example of which would be the popular log sum metric - is applicable to a particular demand problem, namely the probabilistic demand for a discrete good in the absence of income effects. In particular, S&R’s welfare measure is not applicable to demand problems characterised by non-zero income effects.
Second, having adopted S&R as the basis for conducting welfare analysis of a given discrete choice problem, it is necessary to impose four \textit{a priori} assumptions on the specification of deterministic utility:

- **Assumption I:** for each alternative, equivalence (in absolute terms) between the conditional marginal utilities of income and price, i.e. \(-\partial W_i / \partial p_1 = \partial W_i / \partial y, -\partial W_2 / \partial p_2 = \partial W_2 / \partial y\)

- **Assumption II:** common conditional marginal utility of income across alternatives, i.e. \(\partial W_i / \partial y = \partial W_2 / \partial y\)

- **Assumption III:** common conditional marginal utility of price across alternatives, i.e. \(\partial W_i / \partial p_1 = \partial W_2 / \partial p_2\)

- **Assumption IV:** the conditional marginal utility of income, \(\lambda\), is independent of the prices of goods 1 and 2, i.e. \(\partial \lambda / \partial p_1 = \partial \lambda / \partial p_2 = 0\)

Should any of these assumptions be relaxed, then the model specification would be incompatible with S&R’s welfare measure (45).

Third, for most practical welfare analyses, policy interventions would be expected to influence prices in several related markets, making it appropriate to ensure that measures of welfare are path independent. It is clear that S&R’s welfare measure (45), a practical example of which is the popular log sum measure, embodies this property of path dependence. However, an important question that remains to be answered is whether path independence is motivated by theoretical or empirical considerations. In other words, does theory support the proposition that the probabilistic demand could be subject to an income effect (of a price change)? If it does, then S&R’s welfare measure, which is essentially Marshallian, would be an appropriate approximation to Hicksian welfare only in empirical contexts that are not subject to an income effect. If theory does not support the proposition of an income effect on the probabilistic demand, then S&R’s welfare measure would be valid in all empirical contexts.
Fourth, if policy interests motivate the specification of different consumer segments, then these segmentations should be applied to the conditional marginal utilities of both income and price, thereby ensuring that Assumptions I-IV hold within each segment.
ACKNOWLEDGEMENTS

We wish to give particular thanks to Robert Cochrane, whose questions and comments on an earlier draft forced us to clarify and strengthen our arguments. Without his input, the paper would have been much inferior. Thanks are also due to Andrew Daly for comments on an earlier draft, and to participants at the KTH-ITS research seminar in Leeds on 17th-18th June 2010. All remaining errors and misunderstandings are of course the responsibility of the authors.
ANNEX: Derivation of the probabilistic demand, following Hau

The purpose of this annex is to relate Hau (1985, 1987) to the present paper. In what follows, we shall summarise Hau’s analysis, but using S&R’s framework, and the same simple problem of two goods. Note that Hau’s approach follows a similar vein to the earlier work of Domencich & McFadden (1974).

Hau derives the demand for the discrete choice alternative as a good in its own right. Broadly speaking, this might be seen as analogous to the restricted case of S&R, where we assume $x_i = x_i^c = 1$ and adopt the notion of the representative consumer. In this case, the aggregate uncompensated (12) and compensated (13) demands simplify to, respectively:

$$X_i = \sum_{i=1}^{N} x_i^i (p_1, p_2, y) = N\pi_1 (p_1, p_2, y)$$

$$X_i^c = \sum_{i=1}^{N} x_i^c (p_1, p_2, u^i) = N\pi_i^c (p_1, p_2, \{u^i\})$$

As before, we assume that $\pi_1 = \pi_i^c$ holds in equilibrium, implying that $X_i = X_i^c$. In passing, it is worth remarking that this presentation has more in common with S&R’s third ‘rationale’ for discrete choice (see the introduction to the present paper), rather than the first rationale, which was the focus of S&R’s own paper.

If we adopt the utility specification (26) then it is appropriate, given the presence of the random error term, to take expectations when aggregating utilities. More specifically, if we assume that the random error is IID Gumbel, then the discrete choice model adopts the logit form, and the expected maximum utility is given by the so-called log sum expression:

$$E(\max\{\bar{v}_i, \bar{v}_2\}) = E(\max\{(W_i + \epsilon_1), (W_2 + \epsilon_2)\}) = \log(\exp(W_i) + \exp(W_2)) \quad (A1)$$

Having taken expectations, this might be interpreted as the maximum utility for a representative consumer. For brevity, let us denote:
\[ E(\tilde{\nu}_k) \equiv E\left(\max\{\tilde{\nu}_1, \tilde{\nu}_2\}\right) \]  

(A2)

where we again index the chosen good \( k \).

Hau applies the log sum (A1) to Roy’s identity, yielding a statement of probabilistic demand. Since Hau does not show his working, let us derive the probabilistic demand for ourselves. This will serve to expose some important features of Hau’s analysis that coincide with S&R’s.

First, let us consider the primal perspective for good 1 (i.e. assuming \( k = 1 \), noting that the same approach would apply analogously to good 2), and begin by partially differentiating the expected maximum utility (A1) with respect to price and income respectively:

\[
\frac{\partial E(\tilde{\nu}_k)}{\partial p_1} = -\frac{\partial E(\tilde{\nu}_k)}{\partial W_1} \frac{\partial W_1}{\partial p_1} 
\]

(A3)

\[
\frac{\partial E(\tilde{\nu}_k)}{\partial y} = \frac{\partial E(\tilde{\nu}_k)}{\partial W_1} \frac{\partial W_1}{\partial y} + \frac{\partial E(\tilde{\nu}_k)}{\partial W_2} \frac{\partial W_2}{\partial y}
\]

(A4)

where:

\[
\frac{\partial E(\tilde{\nu}_k)}{\partial W_1} = \frac{\partial \left(\log(\exp(W_1) + \exp(W_2))\right)}{\partial W_1} = \frac{1}{\exp(W_1) + \exp(W_2)} \exp(W_1) = \pi_1
\]

Note that (A3) implies that \( \partial W_1/\partial p_2 = 0 \), thereby ruling out models of the ‘mother logit’ form (McFadden et al., 1978). Substituting for \( \partial E(\tilde{\nu}_k)/\partial W_1 \) in (A3) and (A4):

\[
\frac{\partial E(\tilde{\nu}_k)}{\partial p_1} = -\pi_1 \frac{\partial W_1}{\partial p_1} 
\]

(A5)

\[
\frac{\partial E(\tilde{\nu}_k)}{\partial y} = \pi_1 \frac{\partial W_1}{\partial y} + \pi_2 \frac{\partial W_2}{\partial y}
\]

(A6)

Combining (A5) and (A6), we can apply Roy’s identity to expected maximum utility, as follows:
\[
\frac{\partial E(\tilde{v}_k)/\partial p_i}{\partial E(\tilde{v}_k)/\partial y} = -\pi_1 \frac{\partial W_i/\partial p_i}{\partial W/\partial y}
\]

As was noted earlier, in the restricted case \( \tilde{x}_1 = \tilde{x}_2 = 1 \), Shephard’s lemma (and by implication Roy’s identity also) yields the probabilistic demand; the outcome of (A7) should therefore be \( \pi_1 \). For purposes of practical implementation, the question arises as to whether this requirement imposes particular restrictions on the specification of the probabilistic demand model. One approach would be to apply Assumption II from section 3.3 (i.e. \( \partial W_i/\partial y = \partial W_2/\partial y = \lambda \)), in which case Roy’s identity simplifies to:

\[
\frac{\partial E(\tilde{v}_k)/\partial p_i}{\partial E(\tilde{v}_k)/\partial y} = -\pi_1 \frac{\partial W_i/\partial p_i}{\partial W/\partial y}
\]

If we also apply Assumption I (i.e. \( -\partial W_i/\partial p_i = \partial W_i/\partial y \) and \( -\partial W_2/\partial p_2 = \partial W_2/\partial y \), noting that Assumptions I and II would together further imply Assumption III), then (A7) yields the choice probability:

\[
\frac{\partial E(\tilde{v}_k)/\partial p_i}{\partial E(\tilde{v}_k)/\partial y} = \pi_1
\]

In fact, Hau arrives at the same result but via the alternative (but arguably less intuitive assumption) that the partial derivative of expected maximum utility with respect to income (i.e. \( \partial E(\tilde{v}_k)/\partial y \)) is equal to the conditional marginal utility of price (i.e. \( \partial W_i/\partial p_i \)). Moreover, irrespective of which rationale is adopted, we can see that, by imposing assumptions on the specification of both random error (i.e. IID Gumbel) and deterministic utility (i.e. Assumptions I and II) within the probabilistic demand model, Roy’s identity yields a statement of demand simply in terms of discrete choice probability. Of course, the same result could arise from
(A7) even in the absence of these assumptions, but that would be a purely empirical matter.

An alternative rationale for arriving at Assumptions I, II and III would be to acknowledge that, according to theory (e.g. Chapter 3 of Johansson, 1987), the following two identities should hold by definition:

\[ \frac{\partial E(\tilde{v}_k)}{\partial p_i} = -\pi_i \frac{\partial W_i}{\partial y} \] (A8)

\[ \frac{\partial E(\tilde{v}_k)}{\partial y} = \frac{\partial W_i}{\partial y} \] (A9)

In order that (A8) gives the same result as (A5), it is clear that Assumption I must apply, and in order that (A9) gives the same result as (A4), Assumption II must apply.

Having considered the primal perspective, let us now proceed to the dual perspective, again looking at good 1. Hau begins by assuming that the expenditure necessary to realise expected maximum utility \( E(\tilde{v}_k) \) is given by the identity:

\[ e(E(\tilde{v}_k)) = \frac{E(\tilde{v}_k)}{\lambda} \] (A10)

where:

\[ \frac{\partial E(\tilde{v}_k)}{\partial y} = \frac{\partial W_i}{\partial y} = \lambda \]

In fact, this is a very strong assumption, since it imposes constant marginal utility of income. That is to say, if prices change with income fixed then the change in expenditure will be proportional to the change in utility. Another way of rationalising this assumption would be to note that it implies the absence of an income effect.

Differentiating with respect to the price of good 1, we have:
\[
\frac{\partial e(E(\tilde{v}_k))}{\partial p_i} = E(\tilde{v}_k) \cdot \frac{\partial \lambda}{\partial p_i} + \lambda \cdot \frac{\partial E(\tilde{v}_k)}{\partial p_i},
\]

If we apply Assumption IV, then the partial derivative of the expected maximum utility with respect to income will be independent of price:

\[
\frac{\partial (\partial E(\tilde{v}_k)/\partial y)}{\partial p_i} = \frac{\partial \lambda}{\partial p_i} = 0
\]

(A11)

Indeed, this assumption goes hand-in-hand with the constant marginal utility of income (A10); each implies the other. Given Assumption I, we can then apply Shephard's lemma, thus:

\[
\frac{\partial e(E(\tilde{v}_k))}{\partial p_i} = \frac{\partial E(\tilde{v}_k)/\partial p_i}{\partial E(\tilde{v}_k)/\partial y} = \pi
\]

An important practical implication of the above derivation is that it implies tight restriction on permissible specifications of deterministic utility. More specifically, the conditional marginal utility of income should be common across alternatives, and probabilistic demand should be invariant to changes in income.
REFERENCES


