ON THE INTEGRABILITY CONDITIONS FOR DISCRETE TRAVEL CHOICE

J Nicolás Ibáñez and Richard Batley

Institute for Transport Studies, University of Leeds, UK

Summary:

We repeat McFadden's (1981) derivation of the integrability conditions for discrete choice, relaxing several restrictions within a more general analysis. This serves to define the subset of discrete choice models which are faithful to integrability.

Abstract:

In establishing the validity of discrete choice models for economic analysis, a fundamental issue is whether or not they adhere to the integrability conditions. These conditions ensure that, for any system of demand functions involving a symmetric negative semi-definite substitution matrix, there necessarily exists an underlying utility function from which the demand functions can be derived. Conventionally, the integrability conditions exploit 'continuous' demand theory, wherein preferences are defined on a continuous commodity space. Indeed the conditions are based on the partial derivatives of Hicksian demand functions with respect to price and income, and thus appeal to smooth and continuous demand functions. Discrete choice models may be seen as special case of continuous demand theory, such that choice is restricted to a finite and exhaustive subset of the commodity space, and this provokes some challenges in translating the conventional integrability conditions.

The paper considers a more general, and therefore flexible, model of consumption, arising from the combination of both discrete and continuous choices. In particular, we apply this model to the dual theorem of demand, establishing the integrability conditions applying separately to the discrete and continuous consumptions. The most significant prior work in this vein is McFadden's (1981), although it is useful to note Bates' (2003, p19) description of McFadden's analysis as *'…path-breaking though relatively inaccessible …'.*

Our own paper seeks to promote deeper understanding of McFadden's analysis by repeating his derivation from first principles, and annotating this derivation with commentary throughout. In performing this derivation we reveal a number of important, and possibly restrictive, properties of McFadden's analysis. We also discover that a number of common practical specifications of discrete choice models fail to comply with McFadden's integrability conditions; the validity of these specifications for economic analysis is thus exposed to challenge. Last but not least, we review Koning & Ridder's (2003), disputing their conclusion that the integrability conditions on discrete choice consumption are less restrictive than the conditions for random utility maximisation.

Keywords:

Random utility maximisation, integrability of demand systems, duality, discrete choice, continuous consumption

1. Introduction

> *'Apples or pears? Guns or butter? Liberty or death? We experience life as a series of choices, large and small…'* (Hargreaves-Heap et al., 1992 pvii)

As Hargreaves-Heap et al. elegantly articulate, choice plays a fundamental within the economics of consumer behaviour. But in what sense are consumers really called upon to make such stark choices? In most microeconomics textbooks one will be confronted with smoothly continuous indifference maps, implying that commodities such as apples and pears, whilst inherently discrete variables, can be continuously exchanged. Whilst a discrete representation may in some instances, such as the exchange between apples and pears, offer a more accurate statement of the choice problem, continuity brings appealing tractability, at the expense of some loss in realism. In other instances, such as an exchange between gunpowder and butter, a continuous indifference map may be an entirely appropriate statement of the problem. The present paper considers a more general, and therefore flexible, representation of consumption, arising from the combination of both discrete and continuous choices. In particular, we apply this general model to the dual theorem of demand, establishing the 'integrability' conditions applying to the discrete and continuous constituents. In so doing, we establish the conditions on the model of discrete and continuous consumption to ensure validity for economic analysis. Whilst interest in the integrability conditions for continuous demand systems has accumulated a vast literature, the analogous literature applying to discrete choice is small by comparison. The present paper attempts to redress this imbalance by focussing particular attention on the contribution to consumption from discrete choice.

The conventional apparatus for representing discrete economic choice is Random Utility Maximisation (RUM), which establishes a probabilistic relation between the utility of each alternative within an offered set, and the choices of a population of individuals from that set. In deriving the integrability conditions for RUM, within our general problem of discrete and continuous consumption, it is appropriate to draw reference to the small but significant literature on the theoretical basis of RUM. Daly & Zachary (1976, subsequently published 1978) give necessary and sufficient conditions for the consistency of a probabilistic choice system with utility maximisation, formalising the earlier ideas of Marschak (1960) and Block & Marschak (1960). In paying due credit, it might be noted that, concurrent with but independent to Daly & Zachary, two further research streams considered similar interests, these appearing as Williams (1977) and McFadden (1978). Whilst these three works differ in their focus, each can be considered to have contributed significant insight to the formalisation of RUM, and its subsequent adoption in mainstream econometric practice.

The aforementioned literature on RUM does not, however, devote specific attention to the integrability conditions; conditions which ensure that, for any system of demand functions involving a symmetric negative semi-definite substitution matrix, there necessarily exists an underlying utility function from which the demand functions can be derived. This omission should perhaps come as no great surprise, given the origins of the integrability conditions in terms of continuous consumption. Indeed these conditions are based on the partial derivatives of Hicksian demand functions with respect to price and income, and thus appeal to smooth and continuous demand functions. Discrete choice may be seen as special case of continuous consumption, such that choice is restricted to a finite and exhaustive subset of the commodity space, and this provokes significant challenges in translating the conventional integrability conditions.

Whilst a number of authors have acknowledged these challenges, few have given them detailed attention. The definitive contribution is perhaps that of McFadden (1981), which formalises concepts introduced in the earlier work of Domencich & McFadden (1975). In a more recent contribution to the literature, Koning & Ridder (2003) exploit McFadden's (1981) notion of a representative agent model to propose a discrete choice analogue to the integrability conditions. They then proceed to consider the relation between these conditions and the Daly & Zachary (1976, 1978), concluding that the Daly & Zachary conditions are the more restrictive.

The contributions of our own paper are as follows:

- First, we replicate McFadden's (1981) analysis of the integrability conditions, annotating this with commentary on its rationale and assumptions.
- Second, we suggest that some of the inherent assumptions are restrictive, and propose possible generalisations.
- Third, we demonstrate that Koning & Ridder's analysis is based on a fundamental misunderstanding, and that their conclusion concerning the relation between the Daly & Zachary and integrability conditions is erroneous.

2. Some preliminary remarks

As introduced above, a primary focus of this paper is to provide a detailed commentary on Section 5.7 of McFadden (1981), which is entitled 'Aggregation of Preferences'. It is worth noting McFadden's motivation, which he explains at the outset, thus:

> *'The author is unaware of any general analogue for RUM of the simple sufficient (integrability) conditions of individual utility theory. A restricted, but useful, result of this sort is obtained when individual preferences have sufficient structure to aggregate to a social (indirect) utility function yielding aggregate demands. In this case the home economist's traditional representative consumer with fractional consumption rates can be assigned the social utility function, justifying this approach as an analytic shortcut consistent with some underlying population of utility maximizers who make discrete choices' (McFadden, 1981, p206).*

Although we are ultimately interested in outcomes for the population, the basic unit of analysis is individual preference, which we introduce formally as follows. Following McFadden (1981), consider an individual economic agent with income y, offering him or her the vector $\mathbf{x}$ of R divisible (or continuous) commodities at constant (positive) prices $\mathbf{r}$, together with a discrete choice among the set of alternatives $\mathbf{B}$, wherein each alternative $i \in \mathbf{B}$ entails a vector of measured attributes $\mathbf{w}_i$ which may be consumed at price q_i. Hence, in contrast to the usual treatment of duality in microeconomics texts (e.g. Deaton & Muellbauer, 1980, Chapter 2) that deal exclusively with continuous consumption, the present analysis considers an extension to include both continuous and discrete choice consumption.

It is useful to consider the sense in which choice might be represented as discrete rather than continuous. McFadden (1981) takes the uncomplicated view that discrete choice is relevant to specific forms of economic decision that are inherently 'discrete'. He cites the particular examples of labour force participation, occupation, educational level, marital status, family size, residential and work location, travel mode, and brands of commodity purchases. Given the focus of the present paper on combining discrete and continuous consumption, it is useful to distinguish McFadden's perspective from Small & Rosen's (1981). Small & Rosen note three possible rationales supporting the representation of choice as discrete, 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 (the authors cite the example of travel mode choice in this case). Third, if the search for the optimal consumption bundle entails a choice between alternative corner solutions, then the problem is reduced to discrete units. Hence, Small & Rosen present discrete choice less as an inherent property, and more as a restriction on continuity.

An important issue, in terms of our own analysis, is the interface between discrete and continuous, which perhaps calls for a third perspective, to add to those of McFadden and Small & Rosen. With reference to discussion at the introduction, one might rationalise the vector **x** as allocations of apples and pears within the weekly grocery order, with **B** denoting the choice between an apple and pear at the margin. Indeed one could, without any loss of generality, simply think of **x** as all consumption other than the marginal choice. This third perspective implies a more intimate relationship between discrete and continuous consumption, with discrete choice represented simply as the margin of continuous consumption.

Returning to the main theme of the paper, and following the conventions of duality, the consumption optimisation problem defined above can be represented - entirely equivalently - from either of two perspectives. The 'primal' (and perhaps more familiar) problem is one of maximising utility subject to a constraint on budget. The alternative 'dual' problem is one of minimising the expenditure required to attain the maximum utility from the primal problem. Given the marriage between continuous and discrete choice consumption, it is appropriate to derive the integrability conditions for two states: first, for the 'conditional' demand, given the choice of a specific alternative $i \in \mathbf{B}$, and second, for the 'unconditional' demand, considering the complete choice set **B**. The subsequent two sections are devoted to the conditional and unconditional demands respectively, and we derive the primal and dual problems for each.

3. Conditional demand

Define $\tilde{U}$ to be the 'direct' utility to the individual from consuming **x** and choosing alternative $i \in \mathbf{B}$. The primal problem is one of maximising this direct utility subject to budget which, acknowledging the conditioning imposed by the choice *i*, can be defined entirely in terms of the vector of continuous consumption **x**:

$$\begin{aligned} \max_{\mathbf{x}} \quad & \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right) \\ s.t. \quad & \mathbf{rx} + q_i \leq y \rightarrow \left(\lambda\right) \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{1}$$

Here we generalise McFadden's (1981) analysis by specifying utility as a function of not only non-price attributes $\mathbf{w}_i$, but also the price q_i of the discrete choice alternative. Whilst unconventional, such a generalisation does have precedent in continuous demand theory (e.g. Kalmen, 1968; Pollak, 1977), usually motivated by an interest in representing consumption situations where price is a strong indicator of quality. For example, First Class rail travel typically costs more than Standard Class rail travel, but with the benefit of distinct advantages in terms of product quality. Note that, as specified, the attributes of other alternatives within the discrete choice set, that is, $\mathbf{w}_j$ with $j \in \mathbf{B}$ and $j \neq i$, do not impinge upon the problem, thus avoiding any complexities in the manner of the 'mother logit' phenomenon (McFadden et al., 1978).

3.1. Demand for continuous consumption

3.1.1. Primal (utility-maximising) problem and Marshallian demand

Proceeding to solve (1) for **x**, we have that the Lagrangian (ℓ) and the set of equations leading to this solution (see section 5) are as follows:

$$\ell\left(\mathbf{x},\lambda \middle| \mathbf{w}_i,i,q_i,y\right)=\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)-\lambda\left(\mathbf{rx}+q_i-y\right)$$

$$\left.\begin{cases}\mathbf{x}\left(\dfrac{\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)}{\partial\mathbf{x}}-\lambda\mathbf{r}\right)=\mathbf{0}\\ \lambda\left(\mathbf{rx}+q_i-y\right)=0\end{cases}\right\}\rightarrow\left(\mathbf{x}^*,\lambda^*\right) \tag{2}$$

The Karush-Kuhn-Tucker (KKT) conditions applying to problem (1), see also section 5, require that:

$$\frac{\partial\tilde{U}\left(\mathbf{x}^*,q_i,\mathbf{w}_i,i\right)}{\partial\mathbf{x}^*}-\lambda^*\mathbf{r}\le\mathbf{0}\ \mathbf{,}\ \lambda^*\ge 0 \tag{3}$$

Hence, if all the components of continuous consumption are positive (i.e. $\mathbf{x}>0$), the equations determining the solution to (1) adopt the following more familiar form:

$$\left.\begin{cases}\dfrac{\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)}{\partial\mathbf{x}}-\lambda\mathbf{r}=\mathbf{0}\\ \mathbf{rx}+q_i-y=0\end{cases}\right\}\rightarrow\left(\mathbf{x}^*,\lambda^*\right)\ \mathbf{,}\ \lambda^*\ge 0 \tag{4}$$

The KKT conditions applying to this problem have proved a source of contention in the transport literature, particularly when supplementing the money budget with a time budget, as is the norm when analysing the value of travel time (e.g. Bates, 1987; Kockelman, 2001; Hess et al., 2005). We therefore explicate the relevant KKT conditions, which require that at $\left(\mathbf{x}^*,\lambda^*\right)$ the following must hold for any component *m* of the continuous consumption bundle:

$$\lambda=\frac{\dfrac{\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)}{\partial x_m}}{r_m}\ge 0,\forall m=\left\{1,\ldots,R\right\} \tag{5}$$

For any components which take a null value at the optimum (i.e., $x_m^*=0$) we would have instead that:

$$\lambda\ge\frac{\dfrac{\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)}{\partial x_m}}{r_m}\ \mathbf{,}\ \lambda\ge 0 \tag{6}$$

Without lack of generality and for clarity of exposition, we assume hereon that all elements of continuous consumption participate in the optimal bundle. Thus, the solution to the utility maximisation problem in (1) yields the Marshallian demand for the continuous consumption, conditioned on alternative *i*, which we shall note as follows:

$$\mathbf{x}_i=\mathbf{g}_i\left(y-q_i,\mathbf{r},q_i,\mathbf{w}_i,i;\tilde{U}\right) \tag{7}$$

Substituting this solution back into the direct utility function, we derive the conditional indirect utility function (CIUF). The function is 'indirect' in the sense that it represents the utility arising from a specific level of budget, whilst the conditioning pertains to the specific discrete alternative $i \in \mathbf{B}$ under consideration, thus:

$$\begin{aligned}\tilde{U}\left(\mathbf{x}^{*},q_i,\mathbf{w}_i,i\right) &= \max_{\mathbf{x}}\left\{\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)\middle|\mathbf{rx} \le y-q_i\right\} \\ &= \tilde{U}\left(\mathbf{g}_i\left(y-q_i,\mathbf{r},q_i,\mathbf{w}_i,i;\tilde{U}\right),\mathbf{w}_i,i\right) \\ &= V_i\left(y-q_i,\mathbf{r},q_i,\mathbf{w}_i\right)\end{aligned} \qquad \textbf{(8)}$$

Proceeding in this way, it can be seen that the CIUF depends on the price q_i of the discrete choice alternative in two ways: first, through the income available for continuous consumption, and second, through the influence of price on direct utility. Not that if the direct utility in (1) is taken to be independent of price q_i, as in McFadden (1981), then the dependence of the CIUF on price and income can only be through available income, that is, $y-q_i$. Thus, models with CIUF that depend non-linearly on prices alone are let as unconsistent with McFadden's derivation. Furthermore, if one admits the following weaker separability of utility than the one asserted by Jara-Díaz & Videla (1998, #2):

$$\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right) = \tilde{U}\left(\mathbf{x},\mathbf{w}_i,i\right) + \tilde{U}\left(q_i,i\right) \qquad \textbf{(9)}$$

then the utility maximisation problem in (1) becomes exactly that considered by McFadden (1981, #5.5), since the term $\tilde{U}\left(q_i,i\right)$ would not impact upon the solution to such a problem.

In solving this problem, one derives an optimal continuous consumption bundle (**x**) that depends on the price of continuous goods (**r**), the discrete choice attributes other than prices ($\mathbf{w}_i$) and the functional form considered for direct utility ($\tilde{U}$). If direct utility is functional on the prices of discrete alternatives, that is, on q_i, then maximum utility could depend on income *y* and q_i separately; otherwise direct utility will only depend on the quantity $y-q_i$. Furthermore, if discrete choice imposes no specific requirements on the functional form of $\tilde{U}$ and the functional form defining the monetary budget in (1), then the CIUF can be defined *V* rather than V_i.

Adding further explanation, let us consider the standard interpretation of the Lagrange multiplier in (1), λ, as representing the marginal utility of income. From the application of the envelope theorem (see section 5) to problem (1), we can state the equivalence:

$$\lambda^{*} = \frac{\partial \tilde{U}\left(\mathbf{x}^{*},q_i,\mathbf{w}_i,i\right)}{\partial\left(y-q_i\right)} - \left.\left(\frac{\partial \tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)}{\partial\left(y-q_i\right)}\right)\right|_{\mathbf{x}=\mathbf{x}^{*}} \ge 0 \qquad \textbf{(10)}$$

In the absence of any dependence of the direct utility function on available income, λ is equivalent to the marginal utility of income, and its non-negative value imposed by the KKT conditions reflects non-satiation in continuous consumption. In the case where all budget is not exhausted ($\mathbf{rx}+q_i<y$), λ will be zero, indicating satiation. Therefore, where available income is not binding the marginal utility of income is zero, implying also that the marginal utility of consumption is zero. In other words, the consumer is fully satiated.

Moreover, as soon as the individual attains some *positive* marginal utility of income from consuming at least one good x_l, i.e. $\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i) / \partial x_l > 0$, then the consumer will spend all his/her income on consumption, which would in turn make the following two problems equivalent:

$$\boxed{\begin{array}{ll} \max\limits_{\mathbf{x}} & \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i) \\ s.t. & \mathbf{rx} + q_i \leq y \\ & \mathbf{x} \geq \mathbf{0} \end{array}} \;,\; \boxed{\begin{array}{ll} \max\limits_{\mathbf{x}} & \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i) \\ s.t. & \mathbf{rx} + q_i = y \\ & \mathbf{x} \geq \mathbf{0} \end{array}}$$

This equivalence perhaps offers support to a fundamental inference of the Theory of Revealed Preference (Samuelson, 1938). That is to say, in the act of revealing preference the individual also reveals his or her budget (another way of saying this would be to assert that the individual always consumes at the limit of budget). Indeed, the above two problems entail the conventional economic property of ‘non-satiation’, i.e. the consumer derives positive marginal utility from consumption, but budget is binding. It follows that utility should never decrease as budget is increased: more money is usually better, and certainly no worse.

For those cases where the direct utility function determining the optimal bundle of continuous consumption depends on available income, the analysis of (10) shows how a negative marginal utility of income is compatible with non-satiation if it is no smaller than the decrease in the perceived direct utility due to a decrease in the price of the associated discrete choice. Exploiting this fact, together with the relation between the Lagrange multiplier and the CIUF, yields a microeconomic basis for negative (non net) marginal utilities of income.

3.1.2. Dual (cost-minimising) problem and Hicksian demand

The counterpart to the utility maximisation problem formalised in (1) and discussed above is the following cost or expenditure minimisation problem:

$$\begin{array}{ll} \min\limits_{\mathbf{x}} & \mathbf{rx} + q_i \\ s.t. & \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i) \geq u \rightarrow (\gamma) \\ & \mathbf{x} \geq \mathbf{0} \end{array} \tag{11}$$

The different elements of this problem are defined as before, with *u* being an imposed minimum level of direct utility attained by the individual when minimising costs. The Lagrangian and associated set of equations to find the optimal vector **x** of continuous consumption are:

$$\ell(\mathbf{x}, \gamma | q_i, \mathbf{w}_i, i, u) = \mathbf{rx} + \gamma\left(u - \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)\right)$$

$$\left\{\begin{array}{c} \mathbf{x}\left(\mathbf{r} - \gamma \dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial \mathbf{x}}\right) = \mathbf{0} \\ \gamma\left(u - \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)\right) = 0 \end{array}\right\} \rightarrow (\mathbf{x}^*, \gamma^*)$$

The Lagrange multiplier γ represents the marginal cost of utility and is non-negative according to the two KKT conditions applying to (11), see again section 5, that is:

$$\mathbf{r} - \gamma^* \frac{\partial \tilde{U}(\mathbf{x}^*, q_i, \mathbf{w}_i, i)}{\partial \mathbf{x}^*} \geq 0 \;,\; \gamma^* \geq 0$$

Considering again, for the sake of clarity, only positive continuous consumption (i.e. $\mathbf{x}>0$), the equations solving (11) and the associated KKT conditions adopt the following more familiar form:

$$\left\{\begin{array}{c} \mathbf{r}-\gamma\dfrac{\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)}{\partial\mathbf{x}}=\mathbf{0} \\ u-\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)=0 \end{array}\right\}\rightarrow\left(\mathbf{x}^*,\gamma^*\right)\ \mathbf{,}\ \gamma^*\geq 0 \qquad \mathbf{(12)}$$

The absence of degenerated solutions to (11) would imply positive price **r** of the continuous consumption **x** and a direct utility that is not independent of this consumption, rendering the following two problems equivalent:

$$\boxed{\begin{array}{ll} \min\limits_{\mathbf{x}} & \mathbf{rx}+q_i \\ s.t. & \tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)=u \\ & \mathbf{x}\geq\mathbf{0} \end{array}}\ ,\ \boxed{\begin{array}{ll} \min\limits_{\mathbf{x}} & \mathbf{rx}+q_i \\ s.t. & \tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)\geq u \\ & \mathbf{x}\geq\mathbf{0} \end{array}}$$

Note that we distinguish between the prices of discrete choice alternatives (**q**) and the prices of continuous goods (**r**), optimising on the quantity of continuous goods (**x**) given the conditioning of discrete choice $i\in\mathbf{B}$. Indeed, the inclusion of q_i within the cost function has no bearing on the solution, that is, (11) is equivalent to:

$$\begin{array}{ll} \min\limits_{\mathbf{x}} & \mathbf{rx} \\ s.t. & \tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)\geq u \\ & \mathbf{x}\geq\mathbf{0} \end{array} \qquad \mathbf{(13)}$$

Analysing the set of Lagrange multipliers from the primal (1) and dual (11) problems, and substituting for $\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)/\partial x_i$, we can state the following equality for the cases where non-satiation applies ($\lambda^*,\gamma^*>0$):

$$\lambda^*=1/\gamma^*$$

We arrive thus at a definition of the Lagrange multiplier from the dual γ (i.e. the marginal cost of utility), as a relation to the Lagrange multiplier λ (i.e. the marginal utility of income) from the primal. Note that the case corresponding to null net marginal utility of income ($\lambda^*=0$) implies a null value for $\partial\tilde{U}\left(\mathbf{x}^*,q_i,\mathbf{w}_i,i\right)/\partial\mathbf{x}^*$, and therefore an infinitum value for the marginal cost of utility. This means that there is no finite increment of cost at the consumption point $\mathbf{x}^*$ that the individual would be willing to pay to compensate for an increase in his/her achieved utility (since utility remains constant around this point).

Taking the analysis of the dual problem (11) forward, we write the cost function:

$$c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)=\mathbf{rh}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)+q_i \qquad \mathbf{(14)}$$

and differentiate to yield the Hicksian demand for continuous consumption conditional on the alternative *i* being chosen to attain a level of utility *u*, a result referred to as Shephard's Lemma. In this way, we arrive at the solution to the cost minimisation problem formulated in (11):

$$\frac{\partial c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)}{\partial\mathbf{r}}\overset{\left\langle\partial\tilde{U}/\partial\mathbf{r}=\mathbf{0}\right\rangle}{=}\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right) \qquad \mathbf{(15)}$$

If direct utility is *not* functional on the price q_i of the discrete choice alternative then the cost and demand functions, in respect of the continuous consumption, would produce an equivalent result to (15). Moreover the price q_i plays no role in determining the Hicksian demand for continuous consumption unless included in the direct utility function or in supplementary constraints to the budget condition.

If the money budget is exhausted at the optimal value for continuous consumption, or stated equivalently, if the marginal utility is positive ($\partial\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)\big/\partial\mathbf{x}>\mathbf{0}$), then taking the vector of continuous consumption as the Hicksian demand will impose an equality between income and the cost function associated with that demand. That is:

$$c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)\Big|_{u=V_i\left(y-q_i,\mathbf{r},q_i,\mathbf{w}_i\right)}=y \tag{16}$$

3.1.3. Primal and dual equivalence

Returning to the CIUF from the primal problem first introduced in (8), we can exploit this relation between income and the cost function, as follows:

$$\begin{aligned}V\left(c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)-q_i,\mathbf{r},q_i,\mathbf{w}_i\right)&=V\left(\mathbf{r}\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right),\mathbf{r},q_i,\mathbf{w}_i\right)\\&=\max_{\mathbf{x}}\left\{\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)\middle|\mathbf{r}\mathbf{x}\le\mathbf{r}\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)\right\}\\&=\tilde{U}\left(\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right),q_i,\mathbf{w}_i,i\right)\\&=u\end{aligned} \tag{17}$$

Discussion thus far has considered the case where money budget is exhausted, marginal utility is positive, and no other constraints (e.g. such as on time, in the case of travel choice) are pending. If we hold *u* constant and take the total differential of the CIUF in (17) with respect to the price **r** of any continuous commodity, then:

$$\frac{dV\left(z,\mathbf{r},q_i,\mathbf{w}_i\right)}{d\mathbf{r}}=\frac{\partial V\left(z,\mathbf{r},q_i,\mathbf{w}_i\right)}{\partial z}\frac{\partial z}{\partial\mathbf{r}}+\frac{\partial V\left(z,\mathbf{r},q_i,\mathbf{w}_i,i;\tilde{U}\right)}{\partial\mathbf{r}}=0 \tag{18}$$

where:

$$z=c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)-q_i=\mathbf{r}\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right) \tag{19}$$

Through definition (19) we isolate the expenditure devoted to the continuous consumption (remembering that all budget is exhausted). This manipulation is more than a mathematical convenience, supporting weak separability of utility into that associated with the continuous consumption and that associated with discrete choice.

For the moment, we employ this expenditure level *z* and (18) and once again apply Shephard's Lemma, deriving the Hicksian demand for continuous consumption:

$$\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)=-\frac{\left.\dfrac{\partial V\left(a,\mathbf{r},q_i,\mathbf{w}_i,i;\tilde{U}\right)}{\partial\mathbf{r}}\right|_{a=z}}{\dfrac{\partial V\left(z,\mathbf{r},q_i,\mathbf{w}_i,i;\tilde{U}\right)}{\partial z}} \tag{20}$$

This result is known as Roy's identity. It holds for any value of the utility u, since (14) holds for any u, and in particular for its maximum value (u^*) as given by the solution to the primal problem (1). In this latter case, it holds additionally that:

$$u^* = \tilde{U}\left(\mathbf{x}^*, q_i, \mathbf{w}_i, i\right) = \tilde{U}\left(\mathbf{g}_i\left(y - q_i, \mathbf{r}, q_i, \mathbf{w}_i, i; \tilde{U}\right), q_i, \mathbf{w}_i, i\right) = V_i\left(y - q_i, \mathbf{r}, q_i, \mathbf{w}_i\right)$$

and:

$$c\left(\mathbf{r}, q_i, \mathbf{w}_i, i, u^*; \tilde{U}\right) = \mathbf{r}\mathbf{h}_i\left(\mathbf{r}, q_i, \mathbf{w}_i, i, u^*; \tilde{U}\right) + q_i = y$$

Which lead to the following duality relation:

$$\boxed{\mathbf{h}_i\left(\mathbf{r}, q_i, \mathbf{w}_i, i, u^*; \tilde{U}\right) = \mathbf{g}_i\left(y - q_i, \mathbf{r}, q_i, \mathbf{w}_i, i; \tilde{U}\right)} \qquad \textbf{(21)}$$

expressing Roy's identity in terms of the Marshallian demand for continuous consumption.

We have now outlined the key elements of the dual theorem of demand, conditional on discrete choice, and can summarise these diagrammatically by means of Figure 1. Starting out with an objective problem of utility maximisation subject to budget, we inverted the problem to one of cost minimisation subject to achieving a given utility, hence deriving the Hicksian demand (for the continuous consumption) function from the conditional cost function via Shephard's Lemma and the Marshallian demand function from the conditional indirect utility function via Roy's identity.

4. Unconditional demand

Having considered in section 3 the demand conditional on choice, let us now consider the unconditional demand, i.e. the complete set of discrete choice alternatives **B**, again distinguishing between demand for the continuous consumption and demand for the discrete choice alternatives.

4.1. Demand for continuous consumption

We begin by defining the unconditional indirect utility function (UIUF), which is given by the maximal CIUF across the discrete choice alternatives $i \in \mathbf{B}$:

$$\begin{aligned} V^*\left(y\boldsymbol{\iota}_S - \mathbf{q}, \mathbf{q}, \mathbf{r}, \mathbf{w}_\mathbf{B}, \mathbf{B}; \tilde{U}\right) &= \max_{i \in \mathbf{B}}\left\{V_i\left(y - q_i, \mathbf{r}, q_i, \mathbf{w}_i, i; \tilde{U}\right)\right\} \\ &= \max_{i \in \mathbf{B}}\left\{\max_{\mathbf{x}}\left\{\tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right) \middle| \mathbf{r}\mathbf{x} + q_i \le y\right\}\right\} \qquad \textbf{(22)} \\ &= V_{i_{u^*}}\left(y - q_{i_{u^*}}, \mathbf{r}, q_{i_{u^*}}, \mathbf{w}_{i_{u^*}}\right) \end{aligned}$$

where S is the total number of available alternatives (i.e. the size of **B)**, $\mathbf{w}_\mathbf{B}$ is a matrix describing the attributes of each alternative, **q** is the price vector associated with the alternatives and $\boldsymbol{\iota}_S$ is a size-S vector of ones. Let us also write the cost function associated with the UIUF:

$$\begin{aligned} c^*\left(\mathbf{r}, \mathbf{q}, \mathbf{w}_\mathbf{B}, \mathbf{B}, u; \tilde{U}\right) &= \min_{i \in \mathbf{B}}\left\{c\left(\mathbf{r}, q_i, \mathbf{w}_i, i, u; \tilde{U}\right)\right\} \\ &= \min_{i \in \mathbf{B}}\left\{\min_{\mathbf{x}}\left\{\mathbf{r}\mathbf{x} + q_i \middle| \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right) \ge u\right\}\right\} \\ &= \mathbf{r}\mathbf{h}_{i_{c^*}}\left(\mathbf{r}, q_{i_{c^*}}, \mathbf{w}_{i_{c^*}}, i_{c^*}, u; \tilde{U}\right) + q_{i_{c^*}} \end{aligned}$$

The minimum cost occur in only one of the discrete choice alternatives $i \in \mathbf{B}$ (in cases where the minimum is not unique, this can be dealt with by arbitrarily choosing one of the cost-minimising solutions). As discussed earlier, fundamental to duality is the principle that, given a single constraint based on money, the alternative chosen from **B** will be that which minimises cost (denoted here i_{c*}), or equivalently, maximises utility (i_{u*}).

Exploiting Shephard's Lemma in a similar manner to before, we can derive the Hicksian demand for continuous consumption, but this time applying to the choice set **B** in general rather than conditional on any specific alternative $i \in \mathbf{B}$:

$$\frac{\partial c^*\left(\mathbf{r},\mathbf{q},\mathbf{w_B},\mathbf{B},u;\tilde{U}\right)}{\partial \mathbf{r}} = \frac{\partial}{\partial \mathbf{r}}\left(\min_{i\in\mathbf{B}}\left\{\mathbf{rh}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)+q_i\right\}\right) = \mathbf{h}_{i_{c*}}\left(\mathbf{r},q_{i_{c*}},\mathbf{w}_{i_{c*}},i_{c*},u;\tilde{U}\right)$$

With this vector of continuous consumption $\mathbf{h}_{i_{c*}}$ we obtain utility *u* at a minimum cost equal to $\mathbf{rh}_{i_{c*}} + q_{i_{c*}}$. Once we establish duality between the direct utility function and the cost function, there is no possibility of achieving a higher utility than *u* at a cost less than this minimum. Hence, given the vector of continuous consumption **x** and a discrete choice alternative $i \in \mathbf{B}$, which together would incur a total expenditure of $\mathbf{rx} + q_i$, it is not possible to generate more utility than that arising when the expenditure is $\mathbf{rh}_{i_{c*}} + q_{i_{c*}}$. In other words, the maximum possible utility is *u*.

Using this fundamental relation of duality and returning to the UIUF, we have that:

$$\begin{aligned} V^*\left(\mathbf{z},\mathbf{q},\mathbf{r},\mathbf{w_B},\mathbf{B};\tilde{U}\right) &= \max_{i\in\mathbf{B}}\left\{\max_{\mathbf{x}}\left\{\tilde{U}\left(\mathbf{x},q_i,\mathbf{w}_i,i\right)\middle|\,\mathbf{rx}\le z_i\right\}\right\} \\ &= \max_{i\in\mathbf{B}}\left\{V_i\left(z_i,\mathbf{r},q_i,\mathbf{w}_i\right)\right\} \\ &= V\left(z_{i_{u*}},\mathbf{r},q_{i_{u*}},\mathbf{w}_{i_{u*}}\right) \\ &= u \end{aligned} \qquad \mathbf{(23)}$$

where, in line with the previous isolation of the expenditure devoted to continuous and discrete consumption, **z** refers to the following *S*-size vector:

$$\mathbf{z} = c^*\left(\mathbf{r},\mathbf{q},\mathbf{w_B},\mathbf{B},u;\tilde{U}\right)\boldsymbol{\iota}_S - \mathbf{q} = \left(\mathbf{rh}_{i_{c*}}\left(\mathbf{r},q_{i_{c*}},\mathbf{w}_{i_{c*}},i_{c*},u;\tilde{U}\right)+q_{i_{c*}}\right)\boldsymbol{\iota}_S - \mathbf{q}$$

Note that the fundamental relation of duality we have just discussed applies to (23) by establishing the coincidence between the alternatives that minimise total minimum expenditure to reach the maximum utility level, and vice versa, the coincidence between the alternatives that maximise utility while keeping the associated total expenditure to a minimum, that is, $i^* = i_{c*} = i_{u*}$.

Taking the total differential of V^* with *u* constant, we can derive the following:

$$\frac{dV^*}{d\mathbf{r}} = \frac{\partial V_{i_{u*}}\left(z_{i_{u*}},\mathbf{r},q_{i_{u*}},\mathbf{w}_{i_{u*}}\right)}{\partial z_{i_{u*}}}\frac{\partial z_{i_{u*}}}{\partial \mathbf{r}} + \frac{\partial V_{i_{u*}}\left(z_{i_{u*}},\mathbf{r},q_{i_{u*}},\mathbf{w}_{i_{u*}}\right)}{\partial \mathbf{r}} = 0$$

And, finally, applying Shephard's Lemma for the alternative $i^* = i_{c*} = i_{u*}$, we arrive at Roy's identity, and the unconditional Marshallian demand for the continuous consumption:

$$\boxed{\mathbf{h}_{i*}\left(\mathbf{r},q_{i*},\mathbf{w}_{i*},i*,u;\tilde{U}\right)=-\frac{\left.\dfrac{\partial V\left(a,\mathbf{r},q_{i*},\mathbf{w}_{i*}\right)}{\partial \mathbf{r}}\right|_{a=z_{i*}}}{\dfrac{\partial V\left(z_{i*},\mathbf{r},q_{i*},\mathbf{w}_{i*}\right)}{\partial z_{i*}}}}$$

4.2. Demand for discrete choice

In deriving the Hicksian demand for discrete choices we distinguish between the case where direct utility is functional on q_i, and the case where it is not. For the latter case, which is consistent with McFadden's (1981) analysis, and building on the results presented in section 3, we can assert that:

$$\frac{\partial c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)}{\partial q_j}=\frac{\partial\left(\mathbf{r}\mathbf{h}_i\left(\mathbf{r},\mathbf{w}_i,i,u;\tilde{U}\right)+q_i\right)}{\partial q_j}=\tau_{ij} \qquad \textbf{(24)}$$

Where τ_{ij} is unitary for $i=j$ and null otherwise. Result (24) follows from the cost minimisation problem as formulated in (13) and noting that if the direct utility does not depend on prices, then $\mathbf{h}_i\left(\mathbf{r},\mathbf{w}_i,i,u;\tilde{U}\right)$ does not depend either.

On the other hand, for the cases where direct utility is a function of prices, we can alternatives assert that (note change in functional form for $\mathbf{h}_i$):

$$\frac{\partial c\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)}{\partial q_j}=\frac{\partial\left(\mathbf{r}\mathbf{h}_i\left(\mathbf{r},q_i,\mathbf{w}_i,i,u;\tilde{U}\right)+q_i\right)}{\partial q_j}=\tau_{ij}C \qquad \textbf{(25)}$$

Taking in this latter case the total differential of the UIUF in (23) with respect to **q** (the price of the discrete choice alternatives), and noting that the attained utility *u* remains constant, , we have that:

$$\frac{dV^*}{d\mathbf{q}}=\frac{dV_{i*}}{d\mathbf{q}}=\frac{\partial V_{i*}\left(z_{i*},\mathbf{r},q_{i*},\mathbf{w}_{i*}\right)}{\partial z_{i*}}\frac{\partial z_{i*}}{\partial \mathbf{q}}+\left.\frac{\partial V_{i*}\left(a,\mathbf{r},q_{i*},\mathbf{w}_{i*}\right)}{\partial \mathbf{q}}\right|_{a=z_{i*}}=\mathbf{0} \qquad \textbf{(26)}$$

Where i^* refers to the alternative with minimum cost and maximum utility ($i^*=i_{c^*}=i_{u^*}$) and where the expenditure associated with alternative i^* is defined as follows:

$$z_{i*}=c^*\left(\mathbf{r},\mathbf{q},\mathbf{w}_{\mathbf{B}},\mathbf{B},u;\tilde{U}\right)-q_{i*}=\mathbf{r}\mathbf{h}_{i*}\left(\mathbf{r},q_{i*},\mathbf{w}_{i*},i*,u;\tilde{U}\right)$$

The *i*-th relation in (26) leads to Roy's Identity, since it would impose that:

$$\mathbf{r}\frac{\partial \mathbf{h}_{i*}\left(\mathbf{r},q_{i*},\mathbf{w}_{i*},i*;\tilde{U}\right)}{\partial q_{i*}}=-\frac{\left.\dfrac{\partial V_{i*}\left(a,\mathbf{r},q_{i*},\mathbf{w}_{i*}\right)}{\partial q_{i*}}\right|_{a=z_{i*}}}{\dfrac{\partial V_{i*}\left(z_{i*},\mathbf{r},q_{i*},\mathbf{w}_{i*}\right)}{\partial z_{i*}}} \qquad \textbf{(27)}$$

However, for the remaining relations in (26), that is, when considering the total differential of utility respect prices of other alternatives than *i*, we produce a null result. This is in effect another explanation of the motivation that precludes in general the possibility of mother logit.

If, as in McFadden (1981), direct utility is not functional on the prices of the discrete choice alternatives, then we cannot conclude anything from the application of (26), at least not when following this rationale of imposing an equality between income and the cost function associated with the Hicksian demand (19), since in this case (26) would lead to a trivial result after considering that:

$$\frac{\partial z_{i*}}{\partial \mathbf{q}} = \frac{\partial\left(\mathbf{rh}_{i*}\left(\mathbf{r},\mathbf{w}_{i*},i*,u;\tilde{U}\right)\right)}{\partial \mathbf{q}} = 0 \text{ , } \left.\frac{\partial V_{i*}\left(a,\mathbf{r},\mathbf{w}_{i*}\right)}{\partial \mathbf{q}}\right|_{a=z_{i*}} = \mathbf{0} \qquad \textbf{(28)}$$

In order to derive Roy's identity for this case, one must instead defer to the envelope theorem, and this is the subject of the following section.

5. Derivation of Roy's Identity and Shephard's Lemma for discrete choice using the envelope theorem

Previous discussion on the primal (utility-maximising) and dual (cost-minimising) optimisation processes giving rise to the Marshallian and Hicksian demands, respectively, has employed the apparatus of Lagrange multipliers and the associated Karush-Kuhn-Tucker (KKT) conditions. An alternative perspective on the same problem that also employs these conditions is offered by the envelope theorem (see, for instance, Samuelson, 1947, and Silberberg, 1974). This considers the sensitivity of the objective functions of the primal and dual problems at their optima, that is, the maximum utility and minimum cost, respect changes in the parameters, as distinct from changes in the variables (consumption), defining such problems.

We start by considering the KKT conditions themselves, which refer to a generalisation of the Lagrange multiplier method to admit inequality constraints. Although these conditions were first fully formalised by Kuhn & Tucker (1951), they are mainly based on the earlier work of Karush (1939). The constrained and not necessarily linear optimisation problem under analysis is as follows:

$$\boxed{\begin{array}{lrl} \max\limits_{\mathbf{x}} & f\left(\mathbf{x},\boldsymbol{\kappa}_f\right) & \\ s.t. & \mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right) \le \mathbf{0} & \rightarrow\left(\boldsymbol{\xi}\right) \\ & \mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right) = \mathbf{0} & \rightarrow\left(\boldsymbol{\varsigma}\right) \\ & \mathbf{x} \ge \mathbf{0} & \end{array}} \qquad \textbf{(29)}$$

The variables in this problem are denoted by $\mathbf{x}$ and the parameters and functions defining it are denoted as $\boldsymbol{\kappa} = \left(\boldsymbol{\kappa}_f,\boldsymbol{\kappa}_{\mathbf{m}},\boldsymbol{\kappa}_{\mathbf{n}},f,\mathbf{m},\mathbf{n}\right)$. The two different sets of constraints are associated, respectively, with vectors of Lagrange multipliers $\boldsymbol{\xi}$ and $\boldsymbol{\varsigma}$.

For a problem such as (29), the KKT conditions impose that:

$$\left.\begin{array}{r} \mathbf{x}\dfrac{\partial \ell\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\middle|\boldsymbol{\kappa}\right)}{\partial \mathbf{x}} = \mathbf{0} \\ \dfrac{\partial \ell\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\middle|\boldsymbol{\kappa}\right)}{\partial \mathbf{x}} \leq \mathbf{0} \\ \boldsymbol{\xi}\mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right) = 0 \\ \boldsymbol{\xi} \geq \mathbf{0} \\ \mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right) \leq \mathbf{0} \\ \mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right) = \mathbf{0} \end{array}\right\} \rightarrow \left(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*\right) \qquad \textbf{(30)}$$

where ℓ is the Lagrangian correspondent to (29) and described as follows :

$$\ell\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\middle|\boldsymbol{\kappa}\right) = f\left(\mathbf{x},\boldsymbol{\kappa}_f\right) - \boldsymbol{\xi}\mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right) - \boldsymbol{\varsigma}\mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right)$$

As is shown below, the aforementioned envelope theorem follows from these KKT conditions and just the application of both the chain rule of derivatives and the properties of the first order optimality conditions (FOC). We start the derivation of the theorem by noting that the value for **x** that solves (29) is exclusively a function of the parameters $\boldsymbol{\kappa}$, and which is denoted here by $\mathbf{x}^*\left(\boldsymbol{\kappa}\right)$ or, equivalently, by $\mathbf{x}^*$. The substitution of this optimal value in the objective function of (29) and the application of the KKT conditions state that:

$$f\left(\mathbf{x}^*\left(\boldsymbol{\kappa}\right),\boldsymbol{\kappa}_f\right) = \ell\left(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*\middle|\boldsymbol{\kappa}\right)$$

The left-hand term is usually referred to as a value function. Additionally, the FOC over the Lagrangian are:

$$\left.\frac{\partial \ell\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\middle|\boldsymbol{\kappa}\right)}{\partial \mathbf{x}}\right|_{\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\right)=\left(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*\right)} = \mathbf{0}$$

Considering these results whilst taking the derivative of the value function with respect to any parameter κ_t that belongs to $\left(\boldsymbol{\kappa}_f,\boldsymbol{\kappa}_{\mathbf{m}},\boldsymbol{\kappa}_{\mathbf{n}}\right)$, we finally have that:

$$\begin{aligned} \frac{\partial f\left(\mathbf{x}^*\left(\boldsymbol{\kappa}\right),\boldsymbol{\kappa}_f\right)}{\partial \kappa_t} &= \frac{\partial \ell\left(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*\middle|\boldsymbol{\kappa}\right)}{\partial \kappa_t} \\ &= \frac{\partial}{\partial \kappa_t}\left(f\left(\mathbf{x}^*,\boldsymbol{\kappa}_f\right) - \boldsymbol{\xi}^*\mathbf{m}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{m}}\right) - \boldsymbol{\varsigma}^*\mathbf{n}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{n}}\right)\right) \\ &= \left.\left(\frac{\partial f\left(\mathbf{x},\boldsymbol{\kappa}_f\right)}{\partial \mathbf{x}} - \boldsymbol{\xi}\frac{\partial \mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right)}{\partial \mathbf{x}} - \boldsymbol{\varsigma}\frac{\partial \mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right)}{\partial \mathbf{x}}\right)\right|_{\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\right)=\left(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*\right)} \frac{\partial \mathbf{x}^*}{\partial \kappa_t} \\ &\quad + \left.\left(\frac{\partial f\left(\mathbf{x},\boldsymbol{\kappa}_f\right)}{\partial \kappa_t} - \boldsymbol{\xi}^*\frac{\partial \mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right)}{\partial \kappa_t} - \boldsymbol{\varsigma}^*\frac{\partial \mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right)}{\partial \kappa_t}\right)\right|_{\mathbf{x}=\mathbf{x}^*} \\ &\quad - \frac{\partial \boldsymbol{\xi}^*}{\partial \kappa_t}\mathbf{m}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{m}}\right) - \frac{\partial \boldsymbol{\varsigma}^*}{\partial \kappa_t}\mathbf{n}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{n}}\right) \\ &= \left.\left(\frac{\partial f\left(\mathbf{x},\boldsymbol{\kappa}_f\right)}{\partial \kappa_t} - \boldsymbol{\xi}^*\frac{\partial \mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right)}{\partial \kappa_t} - \boldsymbol{\varsigma}^*\frac{\partial \mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right)}{\partial \kappa_t}\right)\right|_{\mathbf{x}=\mathbf{x}^*} \end{aligned} \qquad \textbf{(31)}$$

The penultimate relation makes use of the KKT conditions and the FOC in the following way:

$$
\begin{aligned}
\mathbf{n}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{n}}\right) &= \mathbf{0} \\
\mathbf{m}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{m}}\right)\frac{\partial \boldsymbol{\xi}^*}{\partial \kappa_t} &= 0 \\
\left.\frac{\partial \ell\left(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma}\middle|\boldsymbol{\kappa}\right)}{\partial \mathbf{x}}\right|_{(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma})=(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*)} &= \left.\left(\frac{\partial f\left(\mathbf{x},\boldsymbol{\kappa}_f\right)}{\partial \mathbf{x}} - \boldsymbol{\xi}\frac{\partial \mathbf{m}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{m}}\right)}{\partial \mathbf{x}} - \boldsymbol{\varsigma}\frac{\partial \mathbf{n}\left(\mathbf{x},\boldsymbol{\kappa}_{\mathbf{n}}\right)}{\partial \mathbf{x}}\right)\right|_{(\mathbf{x},\boldsymbol{\xi},\boldsymbol{\varsigma})=(\mathbf{x}^*,\boldsymbol{\xi}^*,\boldsymbol{\varsigma}^*)} = \mathbf{0}
\end{aligned}
\tag{32}
$$

The second equality in (32) is based on the following complementarity imposed by KKT ($\boldsymbol{\xi}^*\mathbf{m}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{m}}\right)=0$) and the exclusion of degenerated (corner) solutions such as $\boldsymbol{\xi}^* = \mathbf{m}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{m}}\right) = \mathbf{0}$. Note that in this scenario, if $\mathbf{m}\left(\mathbf{x}^*,\boldsymbol{\kappa}_{\mathbf{m}}\right) \neq 0$ then $\boldsymbol{\xi}^*$ will necessarily have to be null and also the values for the $\boldsymbol{\xi}^*$ associated with an infinitesimal change to κ_t.

5.1. Roy's Identity

Next we apply (31) to our utility maximisation problem of interest, formulated in (1), yielding the CIUF conditioned on each of the alternatives ($i \in \mathbf{B}$) available to the individual. In this case the various elements participating in (31) would be as follows:

$$
\begin{aligned}
\boldsymbol{\kappa} &= \left(\mathbf{r}, q_i, \mathbf{w}_i, y\right) \\
f\left(\mathbf{x}^*(\boldsymbol{\kappa}),\boldsymbol{\kappa}_f\right) &= V_i(\boldsymbol{\kappa}) = \tilde{U}\left(\mathbf{x}^*(\boldsymbol{\kappa}), q_i, \mathbf{w}_i, i\right)
\end{aligned}
\tag{33}
$$

Taking the derivatives respect the prices of all the available alternatives ($\mathbf{q}$) and income we arrive at:

$$
\begin{aligned}
\frac{\partial V_i(\boldsymbol{\kappa})}{\partial \mathbf{q}} &= \left.\frac{\partial \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right)}{\partial \mathbf{q}}\right|_{\mathbf{x}=\mathbf{x}^*(\boldsymbol{\kappa})} - \lambda\boldsymbol{\delta}_i \\
\frac{\partial V_i(\boldsymbol{\kappa})}{\partial y} &= \left.\frac{\partial \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right)}{\partial y}\right|_{\mathbf{x}=\mathbf{x}^*(\boldsymbol{\kappa})} + \lambda = \lambda
\end{aligned}
\tag{34}
$$

where $\boldsymbol{\delta}_i$ is a dichotomous vector with as many components as number of available alternatives (size of **B**). The *i*-th component of such vector is unitary and the rest are null. Notice again the non participation of the price of other alternatives other than q_i. Given our focus on a single constraint applying to budget, we have finally that:

$$
\boldsymbol{\delta}_i = -\frac{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial \mathbf{q}} - \left.\dfrac{\partial \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right)}{\partial \mathbf{q}}\right|_{\mathbf{x}=\mathbf{x}^*(\boldsymbol{\kappa})}}{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial y}}
\tag{35}
$$

Hence we arrive at a statement of Roy's identity applying to a discrete choice subject to a single (budget) constraint.

5.2. Shephard’s Lemma

The derivation of Shephard’s Lemma follows directly from the application of the envelope theorem to the (dual) cost minimisation problem in (11). In reference to the notation just introduced, the elements of (11) are described as follows:

$$\begin{gathered}\boldsymbol{\kappa} = \left(\mathbf{r}, q_i, \mathbf{w}_i, u\right) \\ f\left(\mathbf{x}^*\left(\boldsymbol{\kappa}\right), \boldsymbol{\kappa}_f\right) = c\left(\boldsymbol{\kappa}\right) = \mathbf{r}\mathbf{x}^*\left(\boldsymbol{\kappa}\right) + q_i\end{gathered} \qquad \textbf{(36)}$$

And the application of (31) leads to the Lemma:

$$\frac{\partial c\left(\boldsymbol{\kappa}\right)}{\partial \mathbf{r}} = \mathbf{x}^*\left(\boldsymbol{\kappa}\right) - \gamma^* \left.\frac{\partial \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right)}{\partial \mathbf{r}}\right|_{\mathbf{x}=\mathbf{x}^*(\boldsymbol{\kappa})} = \mathbf{x}^*\left(\boldsymbol{\kappa}\right) \qquad \textbf{(37)}$$

Note importantly that this derivation of Shephard’s Lemma, with a single constraint on the feasible set defined for **x**, would necessarily require that direct utility does not depend upon the prices **r** of the continuous consumption. Such a dependence could however be admitted if there were additional constraints ($\mathbf{m}\left(\mathbf{x}, \boldsymbol{\kappa}_\mathbf{m}\right) \le \mathbf{0}$),e.g. a time constraint, and:

$$\gamma^*\left(\boldsymbol{\kappa}, \boldsymbol{\kappa}_\mathbf{m}, \mathbf{r}\right) \frac{\partial \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i, \mathbf{r}\right)}{\partial \mathbf{r}} = -\lambda^*\left(\boldsymbol{\kappa}, \boldsymbol{\kappa}_\mathbf{m}, \mathbf{r}\right) \frac{\partial \mathbf{m}\left(\mathbf{x}, \boldsymbol{\kappa}_\mathbf{m}\right)}{\partial \mathbf{r}} \qquad \textbf{(38)}$$

5.3. Reconciliation with McFadden’s (1981) derivation

Contrast (35) with McFadden’s (1981) definition of Roy’s identity as applies to the Marshallian demand for the discrete choice alternative, which is given by the following:

$$\delta_{i^*,k} = -\frac{\dfrac{\partial V^*\left(y\boldsymbol{\iota}_S - \mathbf{q}, \mathbf{r}, \mathbf{w}_\mathbf{B}, \mathbf{B}; \tilde{U}\right)}{\partial q_k}}{\dfrac{\partial V^*\left(y\boldsymbol{\iota}_S - \mathbf{q}, \mathbf{r}, \mathbf{w}_\mathbf{B}, \mathbf{B}; \tilde{U}\right)}{\partial y}} = \begin{cases} 1 & \text{if } k = i^* \\ 0 & \text{if } k \neq i^* \end{cases} \qquad \textbf{(39)}$$

Recall that the earlier approach of holding utility *u* constant and taking total differentials of any CIUF, or the UIUF, with respect to **q** proved inconclusive, hence our deference to the envelope theorem.

Our final word on the envelope theorem is to note an important implication that follows: McFadden’s definition of Roy’s identity is valid only if direct utility is not functional either on **q** or in *y*, and therefore would not admit choice contexts where price proxies for quality.

6. Aggregation of preferences and derivation of social surplus

Having carefully considered the principal theoretical elements of our analysis, we are now equipped to develop the focus of the paper, which concerns the validity of discrete choice models for economic appraisal. To this end, we develop discussion around two particular results, both of which pertain to the specification of ‘random error’ and are thus relevant to the notion of probabilistic choice. Motivating our interest is McFadden’s (1976, p365) assertion that *‘[…] a model in which the experimenter draws individuals* *randomly from a population with differing, but fixed, utility functions, and offers each a single choice […] is consistent with the classical postulates of economic rationality*'.

6.1. Restriction on random terms of conditional indirect utility functions

McFadden (1981) introduces the concept of a social utility function as the function over which the application of Roy's Identity yields the same choice probabilities as those obtained from the aggregation of all possible values for the conditional indirect utility function (CIUF) faced by consumers and defined above in (8). We will proceed to show that this result implies restriction on the type of random terms associated with these CIUFs.

Before doing so, however, it is useful to relate the focus of this section to our discussion at the outset concerning the distinction between discrete choice and continuous consumption. In what follows, the notion of a social utility function is specific to the discrete choice, and distinct from the continuous consumption. Moreover, in applying the social utility function to welfare analysis, one restricts attention to the demand for the discrete choice alternatives. This representation would seem defensible where the discrete choice and continuous consumption are distinct entities, but less defensible where there is some relationship between the two. An example of such a relationship would be where the discrete choice considers a car purchase decision, and the continuous consumption bundle includes the demand for fuel. Another example would be the more general issue described in section 2, where the discrete choice consumption manifests as the margin of continuous consumption.

Whilst the social utility function does not, in principle, account for such inter-relations, it is interesting to reconcile the latter example with the common practice in discrete choice analysis of calibrating an estimated model to revealed market shares (e.g. Ortúzar & Willumsen, 2001). Such calibration might be seen as an approximation to the continuous consumption **x**. That is to say, the model, whilst focussing particular attention on the marginal choice between an apple and pear from **B**, seeks to replicate the market shares of apples and pears, where the latter give rise to the vector **x**. If one accepts this rationale, then the usual method of forecasting in discrete choice analysis might be seen as an extrapolation from the marginal choice.

For example, conventional forecasting practice would be to project the change in the probability of choosing an apple or pear at the margin, in response to a change in the relative price of an apple at the margin. Moreover, this provokes important questions in interpreting discrete choice data, and implementing these data in models. Is it safe to assume that a consumer's preferences with respect to a discrete choice, say between an apple and pear, will be constant for each marginal unit? If not, how can it be established which marginal unit is being consumed and, it follows, which marginal unit should be invoked as the basis for forecasting by extrapolation? These questions would not appear straightforward to answer.

Having set the interests of the present discussion in context and, in particular, acknowledged some limitations on its applicability, let us return to the detailed analysis. Thus far we have referred to the behaviour of an arbitrary individual, who reveals his or her preferences by determining a vector of continuous consumption **x** and a discrete choice ($i \in \mathbf{B}$), and who we now characterise by means of the vector of individual-specific attributes **s**.

In the preceding analysis, this behaviour has been modelled as a deterministic process whereby the individual compares the utility that he or she can achieve for every feasible combination (**x**,**B**), where feasibility is determined by the (monetary) budget constraints. Recall that the alternative incurring minimum cost is also the one enjoying maximum utility (i.e., $i^* = i_{c^*} = i_{u^*}$), noting that this fundamental relation requires well-behaved cost and utility functions.

For an individual with characteristics **s**, consuming **x** of the *R* continuous commodities, and facing a choice from the set **B** comprised of *S* discrete alternatives with attributes $\mathbf{w}_{\mathbf{B}} = \left(\mathbf{w}_1,\ldots,\mathbf{w}_S\right)$, we can determine a preference ordering over **B**, such that:

$$\tilde{U}\left(\mathbf{x}_i,\mathbf{w}_{\mathbf{B}},i\right):\chi\times\omega\rightarrow\left[0,1\right] \ , \ \forall\mathbf{x}_i\in\chi \ , \ \forall\mathbf{w}_{\mathbf{B}}\in\omega \ , \ \forall i\in\mathbf{B}$$

where χ and ω represent, respectively, the candidate spaces for continuous consumption and for the characteristics defining the alternatives comprising **B**.

The comparison of the values that the function $\tilde{U}$ takes for each of the available alternatives determines the preference amongst them, and in that sense, it serves the same purpose as the direct utility function introduced in (1).

Now aggregating across the population of individuals, consider the different direct utility functions $\tilde{U}$ that order preferences over the alternatives in **B** given **x** and $\mathbf{w}_{\mathbf{B}}$. As a means of supporting this interest, we define a probability measure over the values that $\tilde{U}$ can take, represented by:

$$\mu\left(\tilde{\mathbf{U}}\left(\mathbf{x}_{\mathbf{B}},\mathbf{w}_{\mathbf{B}},\mathbf{B}\right),\mathbf{s}\right) \ , \ \tilde{\mathbf{U}}\left(\mathbf{x}_{\mathbf{B}},\mathbf{w}_{\mathbf{B}},\mathbf{B}\right)=\left\{\tilde{U}\left(\mathbf{x}_i,\mathbf{w}_{\mathbf{B}},i\right):\forall i\in\mathbf{B}\right\} \tag{40}$$

This measure μ denotes how probable it is that an individual with characteristics **s** follows the specific relation given by $\tilde{\mathbf{U}}\left(\mathbf{x}_{\mathbf{B}},\mathbf{w}_{\mathbf{B}},\mathbf{B}\right)$ when maximising his or her utility subject to a budget constraint. Hence, μ takes account of the distribution of tastes in the population of individuals with characteristics **s**.

One way to interpret μ, and in line with McFadden (1981, p202), is to assert that there exists a distribution function for the *S*-size real vector $\tilde{\mathbf{U}}\left(\mathbf{x},\mathbf{w}_{\mathbf{B}},\mathbf{B}\right)$ that permits the definition of choice probabilities. Thus, for any alternative $i\in\mathbf{B}$:

$$P\left(i\middle|y\boldsymbol{\iota}_S-\mathbf{q},\mathbf{r},\mathbf{w}_{\mathbf{B}},\mathbf{B},\mathbf{s}\right)=\mu\left(\left\{\tilde{\mathbf{U}}\left(\mathbf{x},\mathbf{w}_{\mathbf{B}},\mathbf{B}\right)\in\Re^S\middle|v_i\geq v_j,\forall j\in\mathbf{B}\right\},\mathbf{s}\right) \tag{41}$$

where:

$$v_i=V_i\left(y-q_i,\mathbf{r},\mathbf{w}_i\right)=\tilde{U}\left(\mathbf{g}_i\left(y-q_i,\mathbf{r},\mathbf{w}_i,i;\tilde{U}\right),\mathbf{w}_i,i\right)$$

Recall from section 3 that V_i represents a CIUF and $\mathbf{g}_i$ a Marshallian demand and that the remainder of the parameters continue to be defined as first introduced in that section..

The probability measure incorporates heterogeneity within segments through differences in the set of values for $\tilde{\mathbf{U}}\left(\mathbf{x},\mathbf{w}_{\mathbf{B}},\mathbf{B}\right)$. This notion of distribution of tastes assumes that there is a probability that a particular set of values for $\tilde{\mathbf{U}}$ occurs when a population of individuals choose from the set of discrete choice alternatives. This can be represented by supplementing the function $\tilde{U}$ in the utility maximisation problem defined in (1) with an additive quantity ε_j. Thus for any alternative $i\in\mathbf{B}$, we would have that:

$$\begin{aligned} \max_{\mathbf{x}} \quad & \tilde{U}\left(\mathbf{x}, q_i, \mathbf{w}_i, i\right) + \varepsilon_i\left(y, q_i, \mathbf{w}_i\right) \\ s.t. \quad & \mathbf{r}\mathbf{x} + q_i \leq y \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \quad \textbf{(42)}$$

For the moment we consider that $\boldsymbol{\varepsilon}$ does not depend on the quantity of continuous consumption **x** or its price **r**, and that it does not therefore influence the solution to (42) Hence we arrive at the familiar notion of a Random Utility Model (Marschak, 1960; Block & Marschak, 1960), whereby utility is comprised of two additive components, the first referred to as 'deterministic', and the second referred to as 'random error'.

The random terms ($\boldsymbol{\varepsilon}$) represent the aforementioned distribution of tastes, and in so doing embody the probabilistic content of RUM models. Note that the effect of $\boldsymbol{\varepsilon}$ on the direct and indirect utility functions, that is, on $\tilde{\mathbf{U}}$ and $\mathbf{V}$, respectively, is identical if $\boldsymbol{\varepsilon}$ is not determined by continuous consumption.

Thus, for the segment of the population with common characteristics **s**, we can derive the probability of choosing any $i \in \mathbf{B}$, as follows:

$$\begin{aligned} P\left(i \middle| y\mathbf{I}_S - \mathbf{q}, \mathbf{q}, \mathbf{r}, \mathbf{w}_{\mathbf{B}}, \mathbf{s}\right) &= \mu\left(\left\{\boldsymbol{\varepsilon} \in \Re^S \middle| u_i \geq u_j, \forall j \in \mathbf{B}\right\}, \mathbf{s}\right) \\ u_i &= V_i\left(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i\right) + \varepsilon_i\left(y, q_i, \mathbf{w}_i\right) \end{aligned} \quad \textbf{(43)}$$

Implementing (41) and (43), we assume that μ can be represented by a probability density *f* for any value $\boldsymbol{\varepsilon} \in \Re^S$. This in turn imposes conditions on probabilistic choice systems, as first introduced in Daly & Zachary (1976, 1978), to guarantee a proper and invariant–to-translation-and scale joint distribution for $\tilde{\mathbf{U}} + \boldsymbol{\varepsilon}$, or equivalently for $\mathbf{V} + \boldsymbol{\varepsilon}$, as discussed in Ibáñez (2007).

In the absence of further assumptions we must accept that all individuals with characteristics **s** have the same income *y*. More generally, and omitting indexation of budget by **s**, we can write the probability of choosing $i \in \mathbf{B}$ across the population, thus:

$$P\left(i \middle| \mathbf{r}, \mathbf{q}, \mathbf{w}_{\mathbf{B}}\right) = P\left(i \middle| y\mathbf{I}_S - \mathbf{q}, \mathbf{r}, \mathbf{w}_{\mathbf{B}}, \mathbf{s}\right) \phi\left(y\mathbf{I}_S - \mathbf{q}, \mathbf{s}\right) \quad \textbf{(44)}$$

where ϕ indicates the distribution of characteristics in the population. The aggregation of preferences within segments considers all possible values of $\tilde{\mathbf{U}}$ in (40), deriving choice probabilities by each segment **s**, and the aggregation across the population considers (44). To this end, we return to the derivation of Roy's identity for discrete choice in (35), and consider the indicator for the chosen alternative.

For the case where direct utility is a function of neither income nor the price of the discrete choice alternatives, as in McFadden (1981), the change in the CIUF following from a change in either income *y* or the price of the discrete choice alternative q_i is constant, since income and price act together via the quantity $y - q_i$. Hence, Roy's identity can be written as:.

$$\boldsymbol{\delta}_i = -\frac{\dfrac{\partial V_i\left(\boldsymbol{\kappa}\right)}{\partial \mathbf{q}}}{\dfrac{\partial V_i\left(\boldsymbol{\kappa}\right)}{\partial y}} \quad \textbf{(45)}$$

By contrast, where direct utility is functional on q_i we have:

$$\boldsymbol{\delta}_i = -\frac{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial \mathbf{q}} - \left.\dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial \mathbf{q}}\right|_{\mathbf{x}=\mathbf{x}^*(\boldsymbol{\kappa})}}{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial y}} \tag{46}$$

In this latter case, the outcome depends on the balance of two effects; an increase in q_i leads to a decrease in CIUF, but this could be offset by a positive contribution to CIUF depending on the value and sign of $\partial \tilde{U} / \partial \mathbf{q}$ at the optimal continuous consumption $\mathbf{x}^*(\boldsymbol{\kappa})$.

The indexes for the chosen alternatives $\boldsymbol{\delta}_i, \forall j \in \mathbf{B}$ are of the same form when the random terms are introduced within direct utility, that is, as in (42), and provided the random terms are not functional on the continuous consumption **x**. The random error terms can, by means of the envelope theorem, be applied to determine the final value for the CIUF, thus:

$$\boldsymbol{\delta}_i = -\frac{\dfrac{\partial U_i(\boldsymbol{\kappa})}{\partial \mathbf{q}} - \left.\left(\dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial \mathbf{q}} + \dfrac{\partial \varepsilon_i(y, q_i, \mathbf{w}_i)}{\partial \mathbf{q}}\right)\right|_{\mathbf{x}=\mathbf{x}^*}}{\dfrac{\partial U_i(\boldsymbol{\kappa})}{\partial y} - \left.\left(\dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial y} + \dfrac{\partial \varepsilon_i(y, q_i, \mathbf{w}_i)}{\partial y}\right)\right|_{\mathbf{x}=\mathbf{x}^*}} \tag{47}$$

where:

$$U_i(\boldsymbol{\kappa}) = V_i(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i) + \varepsilon_i(y, q_i, \mathbf{w}_i) \text{ , } \boldsymbol{\kappa} = (y, q_i, \mathbf{r}, \mathbf{w}_i) \tag{48}$$

Note, however, that if the random terms do not depend on **x**, it holds that:

$$\left.\frac{\partial \varepsilon_i(\mathbf{x}, \boldsymbol{\kappa})}{\partial \kappa_t}\right|_{\mathbf{x}=\mathbf{x}^*(\boldsymbol{\kappa})} = \frac{\partial \varepsilon_i(\boldsymbol{\kappa})}{\partial \kappa_t}$$

and Roy's identity then simplifies to:

$$\boldsymbol{\delta}_i = -\frac{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial \mathbf{q}} - \left.\left(\dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial \mathbf{q}}\right)\right|_{\mathbf{x}=\mathbf{x}^*}}{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial y} - \left.\left(\dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial y}\right)\right|_{\mathbf{x}=\mathbf{x}^*}} = \frac{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial \mathbf{q}} - \left.\left(\dfrac{\partial \tilde{U}(\mathbf{x}, q_i, \mathbf{w}_i, i)}{\partial \mathbf{q}}\right)\right|_{\mathbf{x}=\mathbf{x}^*}}{\dfrac{\partial V_i(\boldsymbol{\kappa})}{\partial y}}$$

6.1.1. Expectation of Roy's Identity for different random terms

Developing the analysis further, we now consider the unconditional indirect utility function (UIUF) for consumer *i*, which is simply the replication of the CIUF yielding maximum utility:

$$\begin{aligned} V^* &= \max_{j \in \mathbf{B}} \left\{ V(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i) + \varepsilon_i(y, q_i, \mathbf{w}_i) \right\} \\ &= V(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}; \boldsymbol{\varepsilon}) \\ &= V(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}) + \varepsilon_{i^*}(y, q_{i^*}, \mathbf{w}_{i^*}) \end{aligned} \tag{49}$$

This UIUF supports the application of Roy's identity, since it represents the utility maximisation problem faced by the consumer when assigning an optimal level of continuous consumption given the choice of alternative i^*. If we assume that $\boldsymbol{\varepsilon}$ is independent from **x**, and consider the conditioning of V^* on the vector of random terms $\boldsymbol{\varepsilon}$ (the conditioning arises because the influence of $\boldsymbol{\varepsilon}$ on determining i^*), then we can state Roy's identity for any $i \in \mathbf{B}$:

$$\boldsymbol{\delta}_i \left| \tilde{\mathbf{U}} = \boldsymbol{\delta}_i \right| \boldsymbol{\varepsilon} = -\frac{\dfrac{\partial V(\boldsymbol{\theta})}{\partial \mathbf{q}}}{\dfrac{\partial V(\boldsymbol{\theta})}{\partial y}} = \begin{cases} 1 & if\ i = i^* \\ 0 & if\ i \neq i^* \end{cases}, \quad \boldsymbol{\theta} = \left(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}; \boldsymbol{\varepsilon}\right) \tag{50}$$

Hence even when the random terms added to the deterministic part of V^* are allowed to depend on the income and price of the discrete choice alternatives, Roy's identity continues to apply.

Furthermore, in order to exploit Roy's identity to derive the dichotomous variables $\boldsymbol{\delta}_i$ indicating the chosen alternative (the one with the highest CIUF value), we need to ensure that:

$$\frac{\partial V\left(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}\right)}{\partial q_i} = 0 \ \ for\ any\ \ i \neq i^* \tag{51}$$

This result explains why mother logit formulation does not adhere to the microeconomics analysis here presented.

Thus, where $\boldsymbol{\varepsilon}$ is independent of **x**, the random terms can be dependent on $(y, \mathbf{q})$ and still yield a consistent set of choice indicators $\boldsymbol{\delta}_i$ which, integrated over $\boldsymbol{\varepsilon}|\mathbf{q}$, would yield a properly behaved probabilistic choice system. We denote this derivation of the demand for discrete choice as $E_{\boldsymbol{\varepsilon}}\, \partial / \partial q_j$, as opposed to the $\partial / \partial q_j\, E_{\boldsymbol{\varepsilon}}$ presented next.

6.1.2. Roy's identity over expected utilities

An alternative approach to deriving the choice probabilities is through employment of a social utility function. Note again McFadden's (1981, p208) assertion: '*We shall now seek sufficient conditions on preferences such that a social utility function can be defined, with fractional consumption rates for the discrete alternatives yielding the same probabilistic choice system'*.

Unlike (30), this approach employs a process to calculate the choice probabilities that considers the expected value for the UIUF, apply it to Roy's identity, and thus derive the consumption predicted for each of the discrete choices (which will be given in terms of probabilities).

We start by considering the following aggregation of the UIUF, which is termed a social utility function (SUF):

$$\bar{V} = E_{\boldsymbol{\varepsilon}}\left(V^* \middle| \boldsymbol{\varepsilon}\right) = E_{\boldsymbol{\varepsilon}}\left(\max_{i \in \mathbf{B}} \left\{V_i\left(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i\right) + \varepsilon_i\left(y, q_i, \mathbf{w}_i\right)\right\}\right)$$

The application of Roy's identity to this aggregated UIUF or SUF yields:

$$\boldsymbol{\rho} = -\frac{\dfrac{\partial \bar{V}}{\partial \mathbf{q}}}{\dfrac{\partial \bar{V}}{\partial y}} \qquad \textbf{(52)}$$

6.1.3. Comparison of the previous two approaches

Note that if we take expectations over the random terms to calculate $\bar{V}$ before applying (52), then the random terms will not participate directly within Roy's identity; contrast this with the $E_{\boldsymbol{\varepsilon}} \partial/\partial q_j$ process in (50). In order for $\boldsymbol{\rho}$ in (52) to reproduce the choice probabilities derived earlier by aggregation of the choice indexes across the population, then we would have to ensure that:

$$\frac{\partial E_{\boldsymbol{\varepsilon}}\left(V^* \middle| \boldsymbol{\varepsilon}\right)}{\partial\left(\mathbf{q}, y\right)} = E_{\boldsymbol{\varepsilon}}\left(\frac{\partial\left(\boldsymbol{\delta}_{i^*} \middle| \boldsymbol{\varepsilon}\right)}{\partial\left(\mathbf{q}, y\right)}\right) \qquad \textbf{(53)}$$

This equivalence requires that the random terms are independent of the price of the discrete choice alternatives and income, since:

$$\bar{V} = \int_{\boldsymbol{\varepsilon}} \left(\max_{i \in \mathbf{B}} \left\{ V_i\left(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i\right) + \varepsilon_i\left(y, q_i, \mathbf{w}_i\right) \right\} \right) \phi\left(\boldsymbol{\varepsilon}\right) d\boldsymbol{\varepsilon} \qquad \textbf{(54)}$$

and:

$$\begin{aligned} \frac{\partial \bar{V}}{\partial\left(\mathbf{q}, y\right)} &= \int_{\boldsymbol{\varepsilon}} \frac{\partial\left(\max_{i \in \mathbf{B}} \left\{ V_i\left(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i\right) + \varepsilon_i \right\} \right)}{\partial\left(\mathbf{q}, y\right)} \phi\left(\boldsymbol{\varepsilon}\right) d\boldsymbol{\varepsilon} \\ &\quad + \int_{\boldsymbol{\varepsilon}} \max_{i \in \mathbf{B}} \left\{ V_i\left(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i\right) + \varepsilon_i \right\} \frac{\partial \phi\left(\boldsymbol{\varepsilon}\right)}{\partial\left(\mathbf{q}, y\right)} d\boldsymbol{\varepsilon} \end{aligned} \qquad \textbf{(55)}$$

The second term of the right hand term in (55) is null under the aforementioned independence, proving that the random terms that govern the aggregation of preferences cannot depend on either the price of the discrete choice alternatives or income. Such dependencies can participate only through the deterministic parts of the conditional indirect utility functions. It might furthermore be remarked that if the random terms are not allowed to depend on **q** then this provides a clear rationale for excluding models of the mother logit form. Proceeding, let us accept this independence, which allows us to state the following two equalities:

$$-\frac{\dfrac{\partial \bar{V}}{\partial \mathbf{q}}}{\dfrac{\partial \bar{V}}{\partial y}} = -\frac{\displaystyle\int_{\boldsymbol{\varepsilon}} \frac{\partial V_{i^*}\left(\boldsymbol{\theta}\right)}{\partial \mathbf{q}} \phi\left(\boldsymbol{\varepsilon}\right) d\boldsymbol{\varepsilon}}{\displaystyle\int_{\boldsymbol{\varepsilon}} \frac{\partial V_{i^*}\left(\boldsymbol{\theta}\right)}{\partial y} \phi\left(\boldsymbol{\varepsilon}\right) d\boldsymbol{\varepsilon}} \quad , \quad \int_{\boldsymbol{\varepsilon}} \left(\boldsymbol{\delta}_i \middle| \boldsymbol{\varepsilon}\right) \phi\left(\boldsymbol{\varepsilon}\right) d\boldsymbol{\varepsilon} = -\int_{\boldsymbol{\varepsilon}} \frac{\dfrac{\partial V_{i^*}\left(\boldsymbol{\theta}\right)}{\partial \mathbf{q}}}{\dfrac{\partial V_{i^*}\left(\boldsymbol{\theta}\right)}{\partial y}} \phi\left(\boldsymbol{\varepsilon}\right) d\boldsymbol{\varepsilon} \qquad \textbf{(56)}$$

$$\boldsymbol{\theta} \equiv \left(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}; \boldsymbol{\varepsilon}\right)$$

Thus, in order for the SUF $\bar{V}$ to yield the same probability system as that arising from the aggregation of probabilities across individuals, we must impose the following:

$$\frac{\int_{\boldsymbol{\varepsilon}} \frac{\partial V_{i^*}(\boldsymbol{\theta})}{\partial \mathbf{q}} \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon}}{\int_{\boldsymbol{\varepsilon}} \frac{\partial V_{i^*}(\boldsymbol{\theta})}{\partial y} \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon}} = \int_{\boldsymbol{\varepsilon}} \frac{\frac{\partial V_{i^*}(\boldsymbol{\theta})}{\partial \mathbf{q}}}{\frac{\partial V_{i^*}(\boldsymbol{\theta})}{\partial y}} \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon} \qquad \textbf{(57)}$$

$$\boldsymbol{\theta} \equiv \left(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}; \boldsymbol{\varepsilon}\right)$$

Moreover, (57) defines the set of permissible CIUF, which are those satisfying the following separability by income:

$$V_i\left(y - q_i, q_i, \mathbf{r}, \mathbf{w}_i\right) + \varepsilon_i = W_{i1}(y) + W_{i2}\left(q_i, \mathbf{r}, \mathbf{w}_i\right) \qquad \textbf{(58)}$$

Recall that in working towards (57) we imposed the independence of **ε** on income (*y*) and the price of the discrete choice alternative (**q**), and in this case we have that:

$$\frac{\partial V_{i^*}\left(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}; \boldsymbol{\varepsilon}\right)}{\partial y} = \frac{\partial W_{i^*1}(y)}{\partial y} \qquad \textbf{(59)}$$

We arrive thus at exactly the function defined in McFadden (1981, #5.12), and in so doing explicate the properties applying to the random error terms.

6.2. Conditions to guarantee compatibility with utility maximisation by a representative agent

The results in the previous section regarding the derivation of the choice probabilities, by aggregation of all possible values of the conditional indirect utility functions (CIUF) across the population, have made it necessary to guarantee the existence of a correctly specified probability measure for those values.

One approach to implementing such an aggregation is to add the random terms to the direct utility function and take the expected result of the discrete choice demand derived by Roy's identity over all possible values of the random terms. This enables us to avoid any dependency between the random terms and continuous consumption **x**.

Note that this aggregation of choices requires that the distribution of possible values for the CIUFs across the population is proper, and invariant to scale and translation. The latter requirements render the conditions to ensure that a probability choice system derived from the aggregation of preferences is consistent with (direct) utility maximising behaviour to be equivalent to the conditions on the probability choice system derived for every individual. Importantly, this result runs counter to the conclusion of Koning & Ridder (2003).

Koning & Ridder consider the additive income formulation of McFadden (1981, #5.12), and assume only symmetry and negative semi-definiteness of the Jacobean involving the derivatives of the choice probabilities with respect to the prices of the alternatives. However for all the separable-in-income formulations defined by (58), and in particular for the additive one considered by Koning & Ridder and McFadden, application of equation (57) would impose the following:

$$P\left(\mathbf{B} \middle| y\mathbf{I}_S - \mathbf{q}, \mathbf{q}, \mathbf{r}, \mathbf{w}_{\mathbf{B}}, \mathbf{B}, \mathbf{s}\right) = -\frac{\frac{\partial \bar{V}}{\partial \mathbf{q}}}{\frac{\partial \bar{V}}{\partial y}} = -\int_{\boldsymbol{\varepsilon}} \frac{\frac{\partial V_{i^*}(\boldsymbol{\theta})}{\partial \mathbf{q}}}{\frac{\partial V_{i^*}(\boldsymbol{\theta})}{\partial y}} \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon} \qquad \textbf{(60)}$$

$$\boldsymbol{\theta} \equiv \left(y - q_{i^*}, q_{i^*}, \mathbf{r}, \mathbf{w}_{i^*}; \boldsymbol{\varepsilon}\right)$$

Thus, given a value for $\boldsymbol{\varepsilon}$, the ratio of derivatives in the integrand (Roy's identity) yields a *S*-size vector, the elements of which are null, except for a unitary value in the i^*-th position, i.e. the indicator of the alternative where $V_i\left(y-q_i,q_i,\mathbf{r},\mathbf{w}_i;\boldsymbol{\varepsilon}\right)+\varepsilon_i,\forall i\in\mathbf{B}$ is at a maximum. Covering all possible outcomes for $\boldsymbol{\varepsilon}$ by means of weighting the aforementioned vector by its density ($\phi\left(\boldsymbol{\varepsilon}\right)$), we can finally derive choice probability.

Furthermore, given the rule to define the index i^* as the utility-maximising alternative, and interpreting the probability measure in (40) as a distribution function, choice probability can be equivalently expressed:

$$P\left(i\middle|y\mathbf{1}_S-\mathbf{q},\mathbf{q},\mathbf{r},\mathbf{w}_{\mathbf{B}},\mathbf{B},\mathbf{s}\right)=\int_{\boldsymbol{\varepsilon}\in\boldsymbol{\Phi}_i}\phi\left(\boldsymbol{\varepsilon}\right)d\boldsymbol{\varepsilon}\ ,\ \forall i\in\mathbf{B} \qquad \textbf{(61)}$$

where the domain of integration is given for any $j\in\mathbf{B}$ by:

$$\boldsymbol{\Phi}_i\equiv\left\{\boldsymbol{\varepsilon}\in\Re^S\middle|V_i\left(y-q_i,q_i,\mathbf{r},\mathbf{w}_i\right)+\varepsilon_i\geq V_k\left(y-q_k,q_k,\mathbf{r},\mathbf{w}_k\right)+\varepsilon_k,\forall k\in\mathbf{B}\right\} \qquad \textbf{(62)}$$

This in turn implies that the probabilistic choice system must comply with a certain set of conditions, to ensure principally that the distribution of random terms is proper (Daly & Zachary, 1976, 1978), but also invariant to translation (Daly, 2007) and invariant to scale (Ibáñez, 2007).

Thus, if the conditions for the equivalence between the choice probabilities emanating from aggregation of individual preferences and the probabilities derived from the consumption of a representative agent hold (57), we cannot support Koning & Ridder's (2003) conclusion that the latter yields weaker conditions on those probabilities.

7. Synthesis and conclusions

The aspiration of this paper was to conduct a re-assessment of the theoretical requirements on discrete choice models to ensure validity for economic analysis. Following the conventional economic implementation of discrete choice models, we centred this re-assessment on the Random Utility Model (RUM), which formulates a probabilistic representation of the economic choices of a population of individuals from a discrete set of choice alternatives. In considering the validity of RUM for economic analysis, a fundamental requirement is that the model is compliant with the integrability conditions. These conditions ensure that, for any system of demand functions involving a symmetric negative semi-definite substitution matrix, there necessarily exists an underlying utility function from which the demand functions can be derived.

The usual definition of the integrability conditions is in terms of continuous demand theory, wherein preferences are defined on a continuous commodity space. RUM, by contrast, is defined on a discrete space, and may thus be seen as special case of continuous demand theory. Given the significance of the integrability conditions, and the proliferation of RUM in economic practice, it is perhaps surprising to note that few researchers have sought to explicate the compliance of RUM with integrability. The definitive contribution in this regard is McFadden (1981).

Our own paper sought to promote deeper understanding of McFadden's analysis by repeating his derivation from first principles, and annotating this derivation with commentary throughout. More specifically, we considered the preferences of an individual consumer when faced with a portfolio of continuous consumption alongside discrete choice. We defined on this portfolio an objective problem of utility maximisation subject to budget, before inverting the problem to one of cost minimisation subject to achieving a given utility. We demonstrated the derivation of the Hicksian demand functions from the cost function via Shephard's Lemma and the derivation of the Marshallian demand functions from the indirect utility function via Roy's identity. We presented this analysis from the perspectives of both conditional demand (i.e. for a specific discrete choice alternative) and unconditional demand (i.e. for the complete set of discrete choice alternatives). Three findings are of particular note:

First, in performing this derivation, we revealed a number of important, and possibly restrictive, properties of McFadden's analysis. For example, implicit within McFadden's model is an assumption that direct utility is not functional on either the price of discrete choice alternatives or income. Whilst such an assumption might on first inspection appear uncontroversial (and indeed attractive), it is important to note that this assumption is violated by specifications of RUM routinely employed in practical discrete choice modelling. The implications of the additive income form describing McFadden's model have also been discussed.

Second, a particularly obscure element of McFadden's derivation is his progression from discrete deterministic choice to discrete probabilistic choice, via the mechanism of probability measures. In this regard, we significantly expanded upon the detail of McFadden's derivation, presenting the clearest and fullest account hitherto. This revealed important restrictions on the manner by which utilities can be aggregated across as sample of individual decision-makers; interestingly this issue again follows from the dependence of direct utility on price and income.

Third, on concluding our derivation we reviewed Koning & Ridder's (2003) recent treatment of the integrability conditions for discrete choice, which exploited McFadden's (1981) notion of a representative agent model to propose a discrete choice analogue to the integrability conditions. We demonstrated that Koning & Ridder's treatment is based on a fundamental misunderstanding concerning the necessary and sufficient conditions that give rise to RUM. Indeed, since this misunderstanding has led to practical applications of RUM which might now be challenged on theoretical grounds, our paper offers clarification on the manner in which RUM should be specified in order to ensure consistency with utility maximisation.

8. References

Bates, J.J., 1987, Measuring Travel Time Values with a Discrete Choice Model: A Note, *The Economic Journal*, 97, 386, pp. 493-498.

Bates, J.J., 2003, Economic Evaluation and Transport Modelling: Theory and Practice, *10th International Conference on Travel Behaviour Research*, Lucerne, Switzerland, August.

Block, H.D., and J. Marschak, 1960, Random orderings and stochastic theories of responses, in J. Marschak (ed.), 1974, *Economic Information, Decision and Prediction: Selected Essays* (Volume 1), D. Reidel, Dordrecht, pp.172-217.

Daly, A.J., 2007, Properties of random utility discrete choice models, in B. Heydecker (ed.), *Mathematics in Transport*, Elsevier, Amsterdam, pp.195-210.

Daly, A.J., and S. Zachary, 1976, Improved multiple choice models. *Proceedings of the Fourth PTRC Summer Annual Meeting*, PTRC, London.

Daly, A.J., and S. Zachary, 1978, Improved multiple choice models, in D.A. Hensher and Q. Dalvi (eds.), *Determinants of Travel Choice*, Saxon House, Farnborough, pp. 335-357.

Deaton, A., and J. Muellbauer, 1980, Economics and Consumer Behavior, Cambridge University Press, Cambridge.

Hargreaves-Heap, S., M. Hollis, B. Lyons, R. Sogden, and A. Weale, 1992, The Theory of Choice: A Critical Guide, Blackwell Publishers.

Hess, S., M. Bierlaire, and J.W. Polak, 2005, Estimation of value of travel-time savings using mixed logit models, *Transportation Research Part A: Policy and Practice*, 39(3), pp. 221-236.

Ibáñez, J.N., 2007, On the compatibility of probabilistic choice systems with random utility maximisation. ITS Working Paper 592, Institute for Transport Studies, University of Leeds.

Jara-Díaz, S.R., 2007, Transport Economic Theory, Elsevier Science.

Jara-Díaz, S.R., and J. Videla, 1989, Detection of income effect in mode choice: theory and application, *Transportation Research Part B: Methodological*, 23, 6, pp. 393-400.

Kalman, P.J., 1968, Theory of consumer behavior when prices enter the utility function, *Econometrica*, 36 (3/4), pp.497-510.

Karush, W., 1939, Minima of Functions of Several Variables with Inequalities as Side Conditions, Master's thesis, Department of Mathematics, University of Chicago, Chicago, Illinois.

Kockelman, K.M., 2001, A Model for Time- and Budget-Constrained Activity Demand Analysis, *Transportation Research Part B: Methodological*, 35 (3), pp.255-269.

Koning, R.H., and G. Ridder, 2003, Discrete choice and stochastic utility maximization. *Econometrics Journal*, 6, pp.1-27.

Kuhn, H.W., and A.W. Tucker, 1951, Nonlinear programming, in J. Neyman (ed), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley, pp. 481-492.

McFadden, D., 1976, Quantal choice analysis: a survey, *Annals of Economic and Social Measurement*, 5 (4), pp. 363-390.

McFadden, D., 1981, Econometric models of probabilistic choice, in C. Manski, and D. McFadden (eds.), *Structural Analysis of Discrete Data With Econometric Applications*. The MIT Press, Cambridge, Massachusetts, pp.198-272.

Domencich, T., and D. McFadden, 1975, Urban Travel Demand: A Behavioral Analysis. North-Holland Publishing Co., Amsterdam.

McFadden, D., W. Tye, and K. Train, 1978, An application of diagnostic tests for the Independence from Irrelevant Alternatives property of the multinomial logit model, *Transportation Research Record*, 637, pp.39-46.

Marschak, J., 1960, Binary choice constraints and random utility indicators, in J. Marschak (1974), *Economic Information, Decision and Prediction: Selected Essays* (Volume 1). D. Reidel, Dordrecht, pp.218-239.

Ortúzar, J. de D., and L. Willumsen, 2001, *Modelling Transport*, 3rd Edition. John Wiley, Chichester.

Pollak, R.A., 1977, Price Dependent Preferences, *The American Economic Review*, 67 (2), pp.64-75.

Samuelson, P.A., 1938, A Note on the Pure Theory of Consumer's Behaviour. *Economica*, 5, pp.61-71.

Samuelson, P.A., 1947, Foundations of Economic Analysis, Harvard University Press, Cambridge, Massachusetts.

Silberberg, E.,1974, A Revision of Comparative Statics Methodology in Economics, or, How to Do Comparative Statics on the Back of an Envelope, *Journal of Economic Theory*, 7, pp.159-172.

Figure 1. Diagrammatic representation of the dual theorem of demand, conditional on discrete choice