Which demand systems can be generated by discrete choice?

Armstrong, Mark and Vickers, John

Department of Economics, University of Oxford

April 2015

Which demand systems can be generated by discrete choice?*

Mark Armstrong John Vickers

April 2015

Abstract

We provide a simple necessary and sufficient condition for when a multiproduct demand system can be generated from a discrete choice model with unit demands.

Keywords: Discrete choice, unit demand, multiproduct demand functions.

1 Introduction

In a variety of economic settings the decision problem facing agents is one of discrete choice. For example, in markets for durable goods such as cars or refrigerators, each consumer who makes a purchase typically buys one unit of one of the products on offer (or buys nothing). If v_i is a consumer's valuation for product i and p_i is its price, then the rational consumer will buy the product with the best value for money given her preferences, i.e., the highest $(v_i - p_i)$ if that is positive, and will otherwise buy nothing. By specifying a probability distribution for the vector of valuations within the population of consumers, one can derive aggregate multi-product demand as a function of the vector of prices. Such a demand system necessarily involves products being substitutes, but otherwise appears to permit rich possibilities of behaviour.[1]

*Both authors at All Souls College, University of Oxford. Thanks for helpful comments are due to two referees, as well as to Simon Anderson, Sonia Jaffe, Howard Smith and Glen Weyl. Contact information for corresponding author: john.vickers@economics.ox.ac.uk, tel: +44 (0)1865 279300.

[1] For example, Hotelling (1932, section 2) provides an early example of a discrete choice demand system. This example exhibits Edgeworth's Paradox, in which an increase in the unit cost of a product (as a result of imposing a new tax, say) causes a multiproduct monopolist to reduce all of its prices.

In this paper we investigate which aggregate demand functions have discrete choice micro-foundations. With a single product, *any* (bounded) downward-sloping aggregate demand function can be generated by a population of unit-demand consumers—the demand function can simply be interpreted as the fraction of consumers who are willing to pay the specified price for their unit. With more than one product, though, the answer is less obvious. We show that discrete choice foundations for an aggregate demand system (which is bounded and exhibits the usual Slutsky symmetry property) exist if and only if all mixed partial derivatives (with respect to prices) of the *total* quantity demanded are negative. Thus there is a simple test for whether a given demand system is consistent with discrete choice.

Early contributions to the theory and econometrics of discrete choice are surveyed by McFadden (1980), who developed the modern economics of discrete choice analysis in a variety of applications including choices of education and residential location. Relationships between discrete choice models and demand systems for differentiated products are explored in chapter 3 (and elsewhere) of the classic analysis by Anderson, de Palma and Thisse (1992). In particular, their Theorem 3.1 states necessary and sufficient properties of demand functions that ensure these demands are consistent with discrete choice. Their result presumes that consumers must buy one option or another, so that total demand always sums to one. In most situations of interest, however, consumers have, and use, the option to buy nothing, and we provide a result in the same spirit as Anderson *et al.*, but which allows for this. Indeed, the way that total demand varies with prices is the key to our analysis.

More recently, Jaffe and Weyl (2010) show how a linear demand system cannot be generated from (continuous) discrete choice foundations when there are at least two products and buyers can consume an outside option.[2] Jaffe and Kominers (2012) extend this analysis to show how (continuous) discrete choice cannot induce a demand system where demand for each product is additively separable in its own price. The analysis in the present paper sets those contributions in a wider context.

[2]Strictly speaking, they show that linear demand does not have discrete choice foundations where the valuations are continuously distributed (so a density exists). In section 3.2 we show how linear demand is often consistent with a discrete choice model in which the support of valuations does not have full dimension.

The next section states a preliminary result, which is not specific to discrete choice, that individual product demands can be derived from the total demand function. The main section then derives necessary and sufficient conditions for the total demand function to be consistent with discrete choice, which are then illustrated by way of some applications and extensions.

2 A preliminary result

Suppose there are n products, with associated price vector $p = (p_1, ..., p_n)$, where the aggregate demand for product $i = 1, ..., n$ is given by $q^i(p) \geq 0$. We only consider prices in the non-negative orthant $\mathbb{R}^n_+$, and we assume quasi-linear preferences, so that demand q_i is the derivative of an indirect utility function $CS(p)$: $q^i(p) \equiv -\partial CS(p)/\partial p_i$, where $CS(\cdot)$ is convex and decreasing in p. For simplicity, suppose that demand functions are differentiable, in which case we have Slutsky symmetry:

$$\frac{\partial q^i(p)}{\partial p_j} \equiv \frac{\partial q^j(p)}{\partial p_i} \text{ for } i \neq j \ . \tag{1}$$

Given the demand system $q(p)$, define $Q(p) \equiv \sum_{i=1}^n q^i(p)$ to be the total quantity of all products demanded with the price vector p. We make the innocuous assumptions that $Q(0) > 0$ and that $Q(p) \to 0$ as all prices p_i simultaneously tend to infinity.

A result which is useful in the "sufficiency" part of the following analysis, and perhaps of interest in its own right, is the following.[3]

Lemma 1 *Suppose the demand system satisfies (1). Then the demand for product i, $q^i(p)$, satisfies*

$$q^i(p) = -\int_0^\infty \frac{\partial}{\partial p_i} Q(p_1 + t, ..., p_n + t) dt \ , \tag{2}$$

where $Q \equiv \sum_i q^i$ is total demand.

Proof. We need to show that

$$q^i(p) = -\int_0^\infty \frac{\partial}{\partial p_i} Q(p_1 + t, ..., p_n + t) dt = -\int_0^\infty \sum_{j=1}^n \frac{\partial q^j}{\partial p_i}(p_1 + t, ..., p_n + t) dt \ .$$

[3] Expression (2) remains valid if Q is continuous and piecewise-differentiable. (Typically, demand is not differentiable at choke prices which make a product's demand fall to zero.)

But (1) implies that the right-hand side above is equal to

$$-\int_0^\infty \sum_{j=1}^n \frac{\partial q^i}{\partial p_j}(p_1+t,...,p_n+t)dt = -\int_0^\infty \frac{d}{dt}q^i(p_1+t,...,p_n+t)dt = q^i(p)$$

as required. ■

Lemma 1, which is true regardless of whether demand is consistent with discrete choice, implies that the total demand function $Q(\cdot)$ summarises all information about the demands for individual products, which can be recovered from total demand via the procedure (2).[4]

3 Which demand systems are consistent with discrete choice?

We wish to understand which restrictions on $q(p)$ are implied if this demand system can be generated by the simplest discrete choice model. By "discrete choice model" we mean, first, that any individual consumer wishes to buy a single unit of one product (or to buy nothing). In particular, a consumer gains no extra utility from buying more than one product or from buying more than one unit of a product. Specifically and furthermore[5], the discrete choice model assumes that a consumer has a valuation v_i for a unit of product i (where valuations can be negative), where the vector of valuations $v = (v_1, ..., v_n)$ is drawn from a joint cumulative distribution function (CDF), denoted $G(v)$, and if she makes a purchase she buys the product which offers the greatest net surplus $v_i - p_i$. If she buys nothing she obtains a deterministic payoff of zero.[6] Faced with price vector p, the type-v consumer in this discrete choice problem will therefore

$$\text{choose product } i \text{ if } v_i - p_i \geq \max_{j \neq i}\{0, v_j - p_j\} \ . \quad (3)$$

[4]For instance, if total demand is additively separable in prices, it follows from (2) that demand for a particular product depends only on its own price. If total demand depends only on the sum of prices, so does the demand for each product.

[5]As we discuss and illustrate in section 3.3 there are settings where consumers buy one unit of one product if they buy at all, but where (3) is not satsified (e.g., because of search or transactions costs). Such settings do not come within the discrete choice model as we have defined it.

[6]The following analysis applies equally to the situation where the consumer's outside option, say v_0, is stochastic, and a consumer buys product with the highest value of $(v_i - p_i)$ provided this is above v_0. However, one can just subtract v_0 from each v_i to return to our set-up with a deterministic outside option of zero.

The demand for product i, $q^i(p)$, is then the measure of consumers who satisfy (3). For most of our discussion we suppose that the distribution for v is continuous—i.e., there is a density function $g(\cdot)$ which generates $G(\cdot)$—which ensures that only a measure-zero set of consumers have a "tie" in (3) and the demand system is well-defined and continuous in prices p. (At various points we also discuss situations where the support of valuations does not have full dimensional support in $\mathbb{R}^n$, although in such cases demand is still continuous in p.) With the choice procedure (3) a consumer buys nothing if and only if $v \leq p$, and so the proportion of consumers who buy nothing with price vector p is just $G(p)$. Figure 1 depicts the pattern of demand with two products, where consumers are partitioned into three regions: those who buy product 1, those who buy product 2, and those who buy neither.

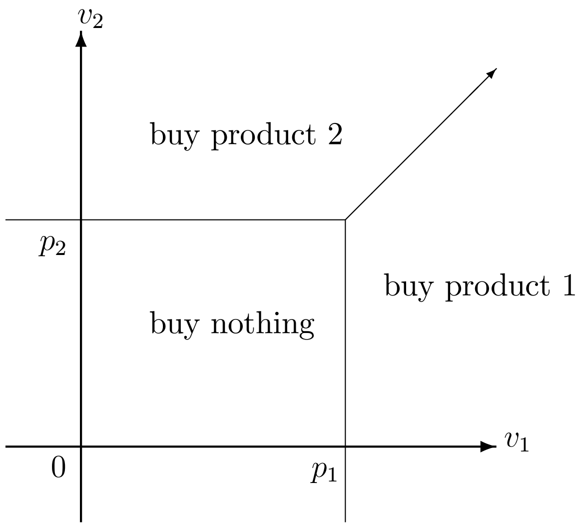

Figure 1: Pattern of demand in discrete choice model

3.1 Necessity

Any demand system arising out of the procedure (3) involves gross substitutes (i.e., cross-price effects are non-negative), since the right-hand side of (3) decreases with p_j. (This can be seen from Figure 1 in the case with two products.) That is to say, a necessary condition for the demand system to be consistent with discrete choice is that $q^i(p)$ weakly increases with p_j for all $j \neq i$.

A second restriction on the demand system $q(\cdot)$ if it is to be consistent with discrete choice is that demand q^i must weakly decrease if all prices increase by the same additive factor. Intuitively, if the price vector increases from $(p_1, ..., p_n)$ to $(p_1 + t, ..., p_n + t)$, no consumer will switch from buying one product to buying another, but some may switch from buying product i to buying nothing. (Again, this is clear from Figure 1.) Regardless of whether it is consistent with discrete choice, as shown in the proof of Lemma 1, any demand system which satisfies (1) satisfies

$$\left.\frac{d}{dt}q^i(p_1 + t, ..., p_n + t)\right|_{t=0} = \frac{\partial}{\partial p_i}Q(p) \ ,$$

where $Q \equiv \sum_j q^j$ is total demand. Therefore, a necessary condition for the demand system to be consistent with discrete choice is that total demand Q weakly decreases with each price p_i.[7]

More generally, for a demand system consistent with discrete choice it must be that total demand Q and the CDF G are related by

$$G(p) \equiv 1 - Q(p) \ , \tag{4}$$

so that $1 - Q$ has the properties of a joint CDF.[8] This is the crucial step in our argument. If a demand system is generated by a discrete choice framework with CDF for valuations G, then $G(p)$, which is the proportion of consumers who buy nothing at price vector p in the discrete choice framework, must be equal to 1 minus the proportion of consumers who buy something, i.e., $1 - Q(p)$. Thus, given any demand system $q(p)$, one can derive the *unique* underlying distribution of valuations which could generate this demand via discrete choice—if such microfoundations are possible—using (4).[9]

[7] Outside the class of demand systems consistent with discrete choice, it is possible to have total demand increase with a price. For example, consider a two-product demand system where $q^i(p_1, p_2) = a_i - b_i p_i + c p_j$. To be consistent with a concave utility function, we require that $b_1 b_2 > c^2$. However, it is still possible that $b_i < c$ for one product, in which case total output $q^1 + q^2$ increases with p_i. More generally, by choosing the units for how products are measured appropriately—by measuring apples in terms of the number of apples and oranges in terms of tons of oranges, say—any demand system with substitutes can be modified so that "total output" increases in a price.

[8] The "1" in (4) simply reflects a normalization of the measure of all consumers to be 1. The analysis could trivially be extended to allow the total measure of consumers to be N, say, in which case total demand Q is bounded by N rather than 1.

[9] More precisely, the CDF for valuations is uniquely determined for $p \geq 0$. As discussed in the proof of Lemma 2, there is some freedom to choose the distribution when some valuations are negative.

If G is a CDF with density function g, then

$$G(p) \equiv \int_{-\infty}^{p_1} \cdots \int_{-\infty}^{p_n} g(v) dv_1 ... dv_n \ . \tag{5}$$

Expression (5) implies that all mixed partial derivatives of G (i.e., which do not involve any "∂p_i" more than once), if they exist, must be non-negative, and the density g can be recovered from G via the partial derivative

$$g(p) \equiv \frac{\partial^n}{\partial p_1 ... \partial p_n} G(p) \ . \tag{6}$$

Since total demand Q satisfies (4), the following necessary conditions on Q are immediate:

Proposition 1 *Suppose that the demand system $q(p)$ is consistent with discrete choice. Then:*
(i) total demand $Q(p) \equiv \sum_{i=1}^{n} q_i(p)$ is continuous at $p = 0$;
(ii) at any price where Q is sufficiently differentiable, for any $1 \leq k \leq n$ and collection of k distinct elements from $\{1, ..., n\}$ denoted $i_1, ..., i_k$ we have

$$\frac{\partial^k}{\partial p_{i_1} ... \partial p_{i_k}} Q(p) \leq 0 \ ,$$

and the corresponding density function for valuations is

$$g(p) = -\frac{\partial^n}{\partial p_1 ... \partial p_n} Q(p) \ .$$

Proposition 1(ii) implies results derived in earlier papers. If $n \geq 2$ then in any region where total demand Q is linear in prices the valuation density must vanish, confirming the result in Jaffe and Weyl (2010). More generally, consider any region where demand for each product is additively separable in its own price, so that $\partial^2 q^i / \partial p_i \partial p_j \equiv 0$ for $j \neq i$. It follows that the full cross-derivative $\partial^n q^i / \partial p_1 ... \partial p_n$ is zero for each demand function q^i, and so the same is true for total demand Q. Again, the density g must vanish in this region, confirming the result derived by Jaffe and Kominers (2012).[10]

[10] A similar argument implies that the density vanishes in any region in which demand functions are additively separable in any non-trivial partition of prices. If $n \geq 2$ and each demand function q^i can be written in the form $A_i(\cdot) + B_i(\cdot)$, where A_i is a function of some non-empty strict subset of prices and B_i is a function of the remaining prices, then again $\frac{\partial^n}{\partial p_1 ... \partial p_n} Q(p) = 0$ and the density vanishes.

However, the fact that the implied density for valuations is zero in all (or almost all) of $\mathbb{R}^n_+$ does not mean that the demand system cannot arise from discrete choice. For instance, in section 3.2 we will see that any (smooth and bounded) demand system without cross-price effects, so that q^i is a function only of its own price p_i, is consistent with discrete choice, although the density will be zero within $\mathbb{R}^n_+$ and no consumer has positive valuation for all n products. We will also see that a linear demand system can be consistent with discrete choice if we allow the support of valuations not to have full dimension in $\mathbb{R}^n$.

Part (i) of Proposition 1 rules out commonly used demand functions which have a discontinuity at $p = 0$. For example, demand which results from homothetic preferences (such as CES preferences) is inconsistent with a discrete choice model. In more detail, suppose the gross utility of the "representative consumer" is homothetic in quantities. It follows that net consumer surplus, $CS(p)$, takes the form $CS(p) = V(P(p))$, where $P(p)$ is a concave and homogeneous degree 1 function of prices and $V(P)$ is a decreasing convex function of the scalar price index P. Then the demand functions are

$$q^i(p) = X(P(p))\frac{\partial P(p)}{\partial p_i} \ , \tag{7}$$

where $X(P) \equiv -V'(P)$. However, the function $\partial P(p)/\partial p_i$ is homogenous degree zero, and such functions cannot be continuous at $p = 0$ (unless they are constant). In sum, any demand system based on a representative consumer with homothetic preferences is not consistent with discrete choice, due to its behaviour when prices are close to zero.[11]

As a final illustration of the use of Proposition 1, consider the demand system whereby total demand takes the form $Q(p_1, p_2, p_3) = (1-p_1)(1-p_2)(1-p_3)$ for prices $0 \le p_i \le 1$.[12] Here, total demand decreases with each price, as required to be consistent with discrete choice, and the corresponding density for valuations from (6) is equal to 1. However, this demand system cannot be consistent with discrete choice since the second cross-partial

[11] In some contexts it is natural to bound prices away from zero, for instance because of production costs. It is then possible for a homothetic demand system to be consistent with discrete choice in this region, although many such demand systems also appear to violate the partial derivative conditions even when bounded away from zero. A two-product CES demand system which is consistent with discrete choice in the region $p_1, p_2 \ge 1$ has $q^i(p_1, p_2) = \sqrt{p_j/(p_i(1+p_1p_2))}$, which takes the form (7) with $P = \sqrt{p_1p_2}$ and $X(P) = 2/\sqrt{1+P^2}$. One can verify that total output $q^1 + q^2$ decreases with each p_i and has a negative cross-partial derivative when $p_1, p_2 \ge 1$. The analysis in section 3.2 then shows this demand system is consistent with discrete choice in this region.

[12] The corresponding individual product demands can be calculated from Lemma 1.

derivatives, $\partial^2 Q/\partial p_i \partial p_j$, are all negative.

Remark on the interpretation of Proposition 1: In the context of discrete choice $1 - Q(p)$ can be interpreted as demand for the outside option of buying nothing—which we may label as notional 'product 0', which by assumption always gives zero consumer surplus—as a function of the prices $p_1, ..., p_n$ of the n actual products. In those terms Proposition 1 is a statement about demand for product 0, and Lemma 1 shows how demand for each product can be derived from demand for product 0. Given that the sum of demands for products 0 to n is by construction equal to one in the discrete choice setting, the method used to derive Lemma 1 also yields that, for any i, demand for each product (including notional product 0) can be derived from demand $q^i(p)$ for product i. In particular, when $\sum_{j=0}^n q^j(p) \equiv 1$, demand for product i can be expressed in terms of the demand function $q^1(p)$ for (say) product $1 \neq i$ by

$$\begin{aligned} q^i(p_0, p_1, ..., p_n) &= -\int_0^\infty \frac{d}{dt} q^i(p_0 + t, p_1, p_2 + t, ..., p_n + t)dt \\ &= \int_0^\infty \frac{\partial}{\partial p_1} q^i(p_0 + t, p_1, p_2 + t, ..., p_n + t)dt \\ &= \int_0^\infty \frac{\partial}{\partial p_i} q^1(p_0 + t, p_1, p_2 + t, ..., p_n + t)dt \ . \end{aligned}$$

The second equality uses the fact that $0 = \sum_{j=0}^n \partial q^j(p)/\partial p_i = \sum_{j=0}^n \partial q^i(p)/\partial p_j$ when demands sum to one. So for any demand system consistent with discrete choice, knowing the demand function for any one product implies the demand functions for all products.

This observation is useful in relating Proposition 1 to Theorem 3.1 of Anderson *et al.* (1992), which was highlighted in the Introduction. For a setting where product demands sum to one, that theorem states, among other things, that consistency with discrete choice requires that all mixed partial derivatives of demand for each product $q^i(p)$ which do not involve its own price p_i be non-negative. Proposition 1 accords with this, but is simpler to state, being just about total demand (equivalently demand for notional product 0) rather than demand for each of n products. Thus it would appear that, with demands by assumption always adding to one, Theorem 3.1 in Anderson *et al.* (1992) could likewise be stated in terms of demand for a single product rather than all.

3.2 Sufficiency

In this section we show, in broad terms, how the necessary conditions outlined in Proposition 1 are also sufficient for the demand system to be consistent with a discrete choice framework. Since we consider only non-negative prices, formula (4) for the candidate CDF for underlying valuations is also defined only on the non-negative orthant $\mathbb{R}^n_+$. Because of this, and since we wish to allow for negative valuations, we need to understand when a function G defined only on $\mathbb{R}^n_+$ can be extended to create a valid CDF defined on the whole space $\mathbb{R}^n$.[13]

Lemma 2 *Suppose G is a sufficiently differentiable function defined on $\mathbb{R}^n_+$ which satisfies $G(0,...,0) = 0$, $G(\infty,...,\infty) = 1$, and for any $1 \le k \le n$ and collection of k distinct elements from $\{1,...,n\}$ denoted $i_1,...,i_k$ we have*

$$\frac{\partial^k}{\partial p_{i_1}...\partial p_{i_k}} G(p) \ge 0 \ . \tag{8}$$

Then $G(\cdot)$ is part of a valid CDF for a continuous distribution on $\mathbb{R}^n$.

Proof. Setting $k = 1$ in (8) implies that G is increasing in each argument, and so G lies in the interval $[0, 1]$ throughout $\mathbb{R}^n_+$. The density g in the region $\mathbb{R}^n_+$ must be given by (6), which from (8) is non-negative.

There are many ways to choose a distribution for v outside $\mathbb{R}^n_+$ which yield the same CDF G when restricted to $\mathbb{R}^n_+$. One way to do so is as follows:
(i) If $v \in \mathbb{R}^n_+$, set $\hat{G}(v) = G(v)$.
(ii) If any component of v is strictly below -1, set $\hat{G}(v) = 0$.
(iii) The remaining case is where v is such that a non-empty subset $S \subset \{1,...,n\}$ of products have valuations in the interval $[-1, 0)$, while remaining products have valuations in $[0,\infty)$. In this case we define

$$\hat{G}(v) = \left(\prod_{i \in S}(1 + v_i)\right) G(v_+) \ , \tag{9}$$

[13] Note that in the following construction the extended density is discontinuous as we cross a plane $v_i \equiv 0$, but that doesn't matter for the argument. One could adjust the argument to make the extended density continuous, if desired.

where v_+ is the vector v with all negative components replaced by zero (i.e., the ith component of v_+ is v_i if $v_i \geq 0$ and 0 otherwise).

One can check that $\hat{G}$ lies in the interval $[0,1]$ throughout $\mathbb{R}^n$, is zero when any v_i is below -1, is continuous throughout $\mathbb{R}^n$, and is weakly increasing throughout $\mathbb{R}^n$. By differentiating (9), one sees that the density corresponding to $\hat{G}$ at a point v such that $k < n$ components of v labelled $i_1, ..., i_k$ are non-negative, while all the remaining components lie in $[-1, 0)$, is

$$\hat{g}(v) = \frac{\partial^n}{\partial p_1 ... \partial p_n} \hat{G}(p) = \frac{\partial^k}{\partial p_{i_1} ... \partial p_{i_k}} G(v_+) \ . \tag{10}$$

From (8), this is non-negative as required.

Define the extended density $\hat{g}$ by (i) $\hat{g}(v) = \frac{\partial^n}{\partial p_1 ... \partial p_n} G(v)$ if each $v_i \geq 0$, (ii) $\hat{g}(v) = 0$ if any $v_i < -1$, and (iii) $\hat{g}(v)$ is given by (10) otherwise. Since $\hat{G}(v) = 0$ if any component $v_i = -1$, it follows that

$$\hat{G}(p) = \int_{-1}^{p_1} \cdots \int_{-1}^{p_n} \frac{\partial^n}{\partial p_1 ... \partial p_n} \hat{G}(p) dv_1 ... dv_n = \int_{-1}^{p_1} \cdots \int_{-1}^{p_n} \hat{g}(v) dv_1 ... dv_n \ .$$

In particular for $p \in \mathbb{R}^n_+$ we have

$$G(p) = \int_{-1}^{p_1} \cdots \int_{-1}^{p_n} \hat{g}(v) dv_1 ... dv_n \ ,$$

and so G defined on $\mathbb{R}^n_+$ is indeed part of a valid CDF. (In particular, the extended density $\hat{g}$ integrates to 1.) ■

Now consider a demand system $q(p)$ which satisfies the required Slutsky symmetry condition (1) such that total demand Q is differentiable throughout $\mathbb{R}^n_+$. It follows that Q is bounded in the neighborhood of $p = 0$, and without loss of generality we can therefore normalize demand so that $Q(0) = 1$. Suppose that $G(p) \equiv 1 - Q(p)$ satisfies the conditions in Lemma 2, i.e., that all the mixed partial derivatives of Q are non-positive. It follows that G is part of a valid CDF for valuations v. By construction, the total demand function which results from the discrete choice model with CDF G is precisely Q. Because the two demand systems—our original $q(p)$ and the demand system implemented by the discrete choice model with CDF G—have the same total demand, Lemma 1 implies that the two demand systems are the same. In particular, $q(p)$ has discrete choice micro-foundations.

We summarise this discussion in the following:

Proposition 2 *Suppose $q(p)$ is a demand system which satisfies (1) such that total demand $Q(p) \equiv \sum_{i=1}^{n} q^i(p)$ is sufficiently differentiable throughout $\mathbb{R}^n_+$, and for any $1 \leq k \leq n$ and collection of k distinct elements from $\{1, ..., n\}$ denoted $i_1, ..., i_k$ we have*

$$\frac{\partial^k}{\partial p_{i_1} ... \partial p_{i_k}} Q(p) \leq 0 \ .$$

Then this demand system can be generated by discrete choice.

A demand system which satisfies the conditions for Proposition 2 must therefore involve gross substitutes, since demands from a discrete choice model do so. This can be seen directly as follows. Since Q is differentiable, we can differentiate both sides of (2) with respect to p_j, where $j \neq i$. This implies that a condition which ensures $\partial q^i / \partial p_j \geq 0$ is that total demand satisfies $\partial^2 Q / \partial p_i \partial p_j \leq 0$ as required by Proposition 2.

Note that any smooth demand system which has no cross-price effects satisfies the conditions of Proposition 2, although the corresponding density g is zero throughout the positive orthant $\mathbb{R}^n_+$. The construction used in Lemma 2 finds a density for valuations which is only positive if only one valuation v_i is positive. To illustrate, suppose there are two products with independent linear demand functions $q^i(p_i) = \frac{1}{2}(1 - p_i)$ (and $q^i = 0$ if $p_i \geq 1$). Then one can check this demand system results from a discrete choice model with density $g(v) = \frac{1}{2}$ if $0 \leq v_i \leq 1$ and $-1 \leq v_j \leq 0$ and $j \neq i$ (and $g(v) = 0$ otherwise).

Proposition 2 applies to demand systems which are differentiable throughout $\mathbb{R}^n_+$, and characterized valid total demand functions in terms of the mixed partial derivatives. This result applies most naturally to demand systems where demand is positive throughout $\mathbb{R}^n_+$. However, the more fundamental property is that the total demand function Q is such that $1 - Q$ is a valid, but not necessarily differentiable, CDF. In the two-variable case, the condition for $G(v_1, v_2)$ for to be a valid CDF is that it is weakly increasing in v_1 and v_2 and the difference $G(v_1^H, v_2) - G(v_1^L, v_2)$ is weakly increasing in v_2 (where $v_1^H > v_2^L$), so that G is increasing and supermodular, i.e., that $Q = 1 - G$ is decreasing and submodular.

To illustrate this more general case, consider the continuous and piecewise-linear demand system depicted on Figure 2.[14] Total demand can be calculated and the candidate

[14] This demand system corresponds to a representative consumer with quadratic gross utility given by $u(q_1, q_2) = \frac{3}{4}(q_1 + q_2) - \frac{3}{8}(q_1^2 + q_2^2 + q_1 q_2)$.

CDF $G = 1 - Q$ then derived, as shown on the figure. One can check that this G is increasing and supermodular, and so this demand system is consistent with discrete choice. Indeed, the required distribution of valuations is that v is equally likely to lie on any of the four bold line segments which make up the boundary of the "kite" shape on the figure, and on any line segment valuations are uniformly distributed.

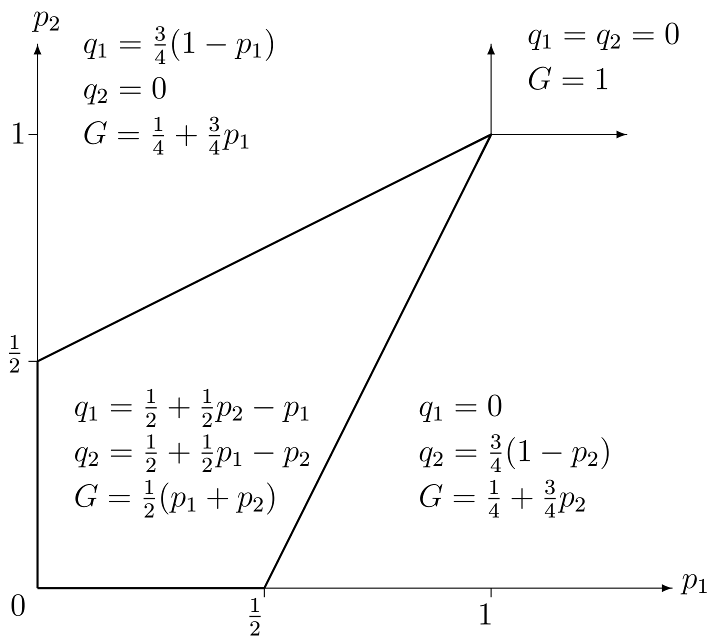

Figure 2: A linear demand system

Conversely, when total demand is not always differentiable but otherwise satisfies the conditions of Proposition 2, the demand system may not be consistent with discrete choice, and so it is not enough just to check locally that the relevant partial derivatives of Q are non-positive. To illustrate, consider the two-product example where $Q(p_1, p_2) = 1 - p_1 p_2$ if $p_1 p_2 \leq 1$ (and otherwise $Q = 0$).[15] The function Q is weakly decreasing in both prices and satisfies $\partial^2 Q / \partial p_1 \partial p_2 = -1$ when $Q > 0$. If this demand system was consistent with discrete choice, the associated CDF for valuations would have to be $G(p_1, p_2) = \min\{p_1 p_2, 1\}$. However, this G is not a valid CDF (unless prices are restricted to lie in the square $[0, 1]^2$), since it does not satisfy the increasing differences property (e.g., here $G(2, p_2) - G(1, p_2)$ decreases with p_2 in the range $\frac{1}{2} \leq p_2 \leq 1$).

[15] Again, Lemma 1 can be used to generate the corresponding individual product demand functions.

4 Applications and extensions

We now consider some examples and extensions of the discrete choice model, and related examples that do not accord with it.

Total demand is a completely monotonic function of an additively separable function of prices: A rich class of demand systems consistent with discrete choice has $1 - Q(p) = Z(\sigma(p))$ as a completely monotonic function of a sum $\sigma(p) \equiv \sum_{i=1}^{n} \alpha_i(p_i)$ of positive, decreasing functions of price, one for each product.[16] (A function $Z : (0, \infty) \to \mathbb{R}$ is said to be *completely monotonic* if for all k the k^{th} derivative, denoted $Z^{(k)}$, has the sign of $(-1)^k$. For our purposes it suffices that this condition holds for $k \leq n$.) Then

$$\frac{\partial^k}{\partial p_1 ... \partial p_k} Q(p) = -Z^{(k)}(\sigma) \prod_{i=1}^{k} \alpha_i'(p_i) < 0 \tag{11}$$

because $Z^{(k)}$ and the product of the α_i' terms both have the sign of $(-1)^k$ so (11) is negative. Proposition 2 then implies that a distribution of valuations can be found which generates this total demand via discrete choice.

The Logit demand system, perhaps the most familiar model of discrete choice, belongs to this class.[17] This demand system has

$$q^i(p) = \frac{1+n}{n} \cdot \frac{e^{-p_i/\mu}}{1 + \sum_j e^{-p_j/\mu}}$$

for some parameter $\mu > 0$. (Demands are normalized by the factor $\frac{1+n}{n}$ to satisfy our convention that $Q = 1$ when $p = 0$.) Here, $1 - Q(p) = Z(\sigma(p))$, where $Z(\sigma) = 1 - \frac{1+n}{n} \frac{\sigma}{1+\sigma}$ and $\sigma(p) = \sum_i e^{-p_i/\mu}$. Also in this class is the case of discrete choice where valuations v_i are independently distributed and non-negative. With $G_i(v_i)$ as the CDF of v_i we can write $1 - Q(p) = \prod_{i=1}^{n} G_i(p_i)$ as $Z(\sigma(p))$, where $Z(\sigma) = e^{-\sigma}$ and $\sigma(p) = -\sum_{i=1}^{n} \log G_i(p_i)$. (In either case, one can check that $Z(\sigma)$ is completely monotonic.)

[16] So that $Q(0) = 1$ and $Q(\infty) = 0$, suppose that each α_i satisfies $\alpha_i(\infty) = 0$, while $Z(\Sigma\alpha_i(0)) = 0$ and $Z(0) = 1$.

[17] See, for example, Anderson *et al.* (1992, section 7.4). The usual micro-foundations for this demand system has consumer valuations—including the value of the outside option—being independent extreme value variables. In particular, the value of the outside option is stochastic. Anderson *et al.* (1992, section 7.4) also present the demand system when product valuations are independent extreme value variables but the outside option has a deterministic value of zero, but this is algebraically messier.

Completely monotonic functions can be used more generally to extend a given discrete choice model to a wider family. For if $Q(p)$ satisfies the conditions of Proposition 2, then so does $\hat{Q}(p) = 1 - \zeta(Q(p))$ where ζ is a completely monotonic function with $\zeta(0) = 1$ and $\zeta(1) = 0$. (One can check that any k^{th} order mixed partial derivative of $\hat{Q}(p)$ is a sum of negative terms.)

Consumer search: We have defined the discrete choice model by condition (3) that the consumer will buy the product with the highest $v_i - p_i \geq 0$, which accords with consumers being able to learn their valuations costlessly. However, the discrete choice model can be used also to analyze some (but not all) settings with search costs, as the following two-product example illustrates. Suppose that the valuation for product i has independent CDF $G_i(v_i)$ (where both valuations are always non-negative), that the consumer knows both prices and can observe v_1 costlessly but that she has to pay search cost s, with independent CDF $F(s)$, to learn v_2.

Assume first that there is free recall, so that the consumer can costlessly return to buy product 1 if she investigates but doesn't end up buying product 2. In this case, the consumer will buy nothing if both (a) $v_1 < p_1$ and (b) either $v_2 < p_2$ or $s > V(p_2) \equiv \int_{p_2}^{\infty}(v_2 - p_2)dG_2(v_2)$. Therefore the proportion who buy nothing is

$$1 - Q(p_1, p_2) = G_1(p_1)[1 - F(V(p_2))(1 - G_2(p_2))] \ . \tag{12}$$

Denoting the square-bracketed term in (12) by $\tilde{G}_2(p_2)$, we have $\tilde{G}_2'(p_2) > 0$ and $1 - Q(p) = G_1(p_1)\tilde{G}_2(p_2)$ satisfies the conditions of the discrete choice model.[18] In short, this model with search costs has a counterpart without them that is consistent with discrete choice.

But that is not the case with costly recall. Suppose that a consumer who investigates product 2 must pay a positive search cost to revisit product 1. Then $Q = 1$ when $p_1 = 0$ and $p_2 = \infty$ because all consumers buy product 1 without searching further. But $Q < 1$ when $p_1 = 0$ and $p_2 > 0$ but is small enough that $F(V(p_2)) > 0$. This is because some consumers with low s and low v_1 will investigate product 2 only to find that $v_2 < p_2$, and when v_1 is below the re-visiting cost they will not wish to return to buy product 1 either.

[18] Since the joint CDF for the valuations in the corresponding discrete choice model is $G(p_1, p_2) = G_1(p_1)\tilde{G}_2(p_2)$, the distribution for v_1 is unchanged from the search model, while the distribution of the valuation for the second product is shifted downwards (since $\tilde{G}_2 \geq G_2$), reflecting the cost needed to discover this valuation.

Therefore, Q is not monotonic in p_2 and the discrete choice model does not apply. To illustrate most starkly, suppose there are no search costs but the consumer cannot return to buy product 1 if she does not purchase it immediately. She will then buy product 1 if $v_1 - p_1 \geq V(p_2)$, and otherwise she buys product 2 if $v_2 \geq p_2$. Thus, she buys nothing with probability

$$1 - Q(p_1, p_2) = G_1(p_1 + V(p_2))G_2(p_2) \ .$$

Then Q must sometimes increase with p_2 when $p_1 = 0$, since Q tends to 1 as p_2 becomes large, and this is inconsistent with a discrete choice model.

Extending discrete choice to allow consumers to buy several products: An extension of the standard discrete choice model allows consumers to buy several products, rather than having to choose just one. The question then arises when this extended notion of discrete choice conforms with the basic one described at the start of section 3. To examine this issue briefly, suppose for simplicity there are two products, that v_i is a consumer's valuation for product $i = 1, 2$ on its own, while her valuation for the bundle of both products is $v_1 + v_2 - z$ for some constant $z \geq 0$. Here, z reflects an intrinsic "disutility" from joint consumption, reflecting an assumption that the products are partial substitutes. (The usual model of discrete choice is the limiting case of this when $z \to \infty$.) The pattern of demand given the pair of prices (p_1, p_2) is shown in Figure 3.[19]

If $F(v_1, v_2)$ is the CDF for (v_1, v_2), then total demand with prices (p_1, p_2) is

$$Q = 1 - F(p_1, p_2) + \{1 - F(p_1 + z, \infty) - F(\infty, p_2 + z) + F(p_1 + z, p_2 + z)\} \ .$$

(Here, the term in brackets $\{\cdot\}$ is the fraction of consumers who buy both products.) Then Q decreases with each price p_i, as needed to be consistent with the usual discrete choice model with single-product demand. The cross-partial derivative is

$$\frac{\partial^2 Q}{\partial p_1 \partial p_2} = f(p_1 + z, p_2 + z) - f(p_1, p_2) \ ,$$

where f is the density function for valuations (v_1, v_2). Thus, if the above expression is always negative—which is the case, for instance, if f decreases with (v_1, v_2)—the demand

[19] This figure is taken from Armstrong (2013). Gentzkow (2007) empirically investigates a related discrete choice model in which some consumers purchase two items.

system induced by this extended discrete choice model is consistent with another basic discrete choice model in which consumers buy at most one product.

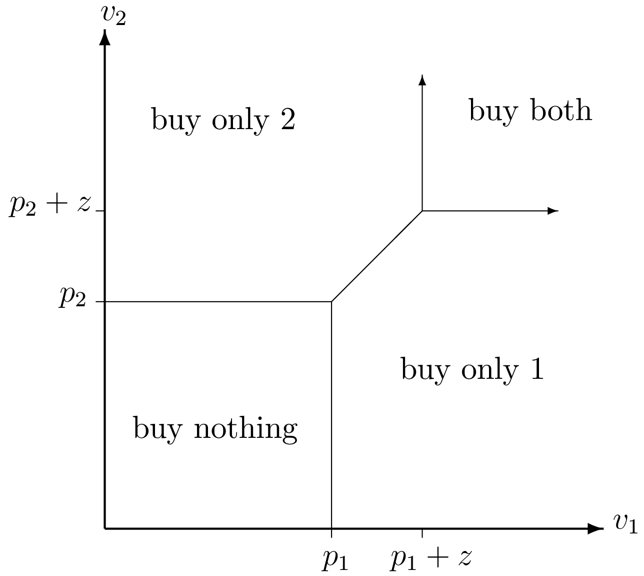

Figure 3: Pattern of demand when products are partial substitutes

Extending discrete choice to allow consumers to buy multiple units of their chosen product: The final extension we examine allows consumers to buy their chosen product in continuous quantities, although as in the basic discrete choice model each consumer buys at most one product.[20] Specifically, suppose that all consumers have the same demand for a given product, and each consumer has demand $x_i(p_i)$ if she buys product i with price p_i. Let $s_i(p_i)$ be the consumer surplus function which corresponds to $x_i(p_i)$. Consumers incur idiosyncratic additive shocks to their surplus vector (e.g., in their "transport costs" to reach a product), denoted $\tau = (\tau_1, ..., \tau_n)$, and the type-$\tau$ consumer chooses the buy the product with the highest value of $s_i(p_i) - \tau_i$ (or buys nothing if $\tau_i \geq s_i(p_i)$ for all products). Let $X^i(s)$ be the fraction of consumers who choose product i when the surplus vector is $s = (s_1, ..., s_n)$.

As in any discrete choice problem of this form, X^i increases with s_i and decreases with

[20] See Hanemann (1984) for an early investigation of this demand model. Anderson *et al.* (1987) show CES demand can be generated from a demand system in which a consumer buys just one product but has downward-sloping demand for the chosen product.

any other s_j. Aggregate demand for product i is

$$q^i(p) = X^i(s(p))x_i(p_i) \ ,$$

and so total demand is

$$Q(p) = \sum_{i=1}^{n} X^i(s(p))x_i(p_i) \ .$$

It follows that

$$\frac{\partial Q}{\partial p_i} = X^i x_i' - x_i \sum_j \frac{\partial X^j}{\partial s_i} x_j \ , \tag{13}$$

which we claim is ambiguous in sign.

To see this, consider the symmetric two-product case with $X^1 = \xi(\delta)$ an increasing function of the surplus difference $\delta \equiv s(p_1) - s(p_2)$ and $X^2 = \xi(-\delta)$. Instances of this are the Hotelling model where $\xi(\delta) = \max\{\frac{1}{2}[1 + \frac{\delta}{t}], 0\}$ where $t < 1$ is a transport cost, and the logit formulation with $\xi(\delta) = \frac{1}{1+e^{-\delta}}$. Suppose that demand for the chosen product takes the exponential form, so that $s(p) = x(p) = e^{-p}$. Then from (13) we have

$$\frac{\partial Q}{\partial p_1} = x_1'\xi - x_1(x_1 - x_2)\xi' = -x_1(\xi + \delta\xi') \tag{14}$$

and

$$\frac{\partial^2 Q}{\partial p_1 \partial p_2} = -x_1 x_2(2\xi' + \delta\xi'') \ . \tag{15}$$

In the Hotelling model with prices such that $|\delta| < t$, (14) implies

$$\frac{\partial Q}{\partial p_1} = -x_1\left(\frac{1}{2} + \frac{\delta}{t}\right) \ .$$

So when $\delta \approx -t$ we have $\frac{\partial Q}{\partial p_1} \approx \frac{x_1}{2} > 0$ and there is inconsistency with the basic discrete choice model. Increasing the high price, p_1, reduces demand from the few consumers who continue to buy that product but causes others to switch to product 2, of which they buy substantially more.

With the logit formulation we have $\xi' = \xi(1 - \xi) > 0$, so from (14)

$$\frac{\partial Q}{\partial p_1} = -x_1[1 + \delta(1 - \xi)]\xi < 0$$

because $\delta > -1$. Expression (15) implies that the cross-partial of Q is

$$\frac{\partial^2 Q}{\partial p_1 \partial p_2} = -x_1 x_2[2 + \delta(1 - 2\xi)]\xi' < 0 \ .$$

Therefore, this demand system with logit shocks to consumer surplus and exponential demand for the chosen product is consistent with the basic discrete choice model.

5 Conclusion

Propositions 1 and 2 together show that, assuming that total demand is differentiable and bounded, the necessary and sufficient condition for consistency with the discrete choice model is that all mixed partial derivatives of total demand be non-positive. (More fundamentally, without requiring differentiability the condition is that $1-Q$ exhibits the required properties of a joint CDF.) This is a strong form of product substitutability.

We have focused on the basic discrete choice model where each consumer buys one unit of one product, specifically the product with highest $(v_i - p_i)$, or else nothing. But in a setting with consumer search and free recall, where consumers do not always buy the product with the highest $(v_i - p_i)$, was shown to be equivalent to a basic discrete choice model that by definition meets the condition. We also showed that situations in which consumers could buy a unit of more than one product, or could buy several units of their chosen product, was sometimes equivalent to the basic unit-choice setting. So the analysis of the basic discrete choice model has more general application.

We have also focused on those situations in which linear prices are used. However, even if an aggregate demand system is consistent with discrete choice with linear prices, it may exhibit very different properties when more ornate tariffs are employed. For instance, when facing unit-demand consumers, a seller can never benefit from the use of two-part tariffs, nonlinear pricing or bundling, while if the seller faced a single consumer with the same aggregate demand it will usually prefer to use a two-part tariff instead of linear prices.

References

Anderson, Simon, André de Palma and Jacques Thisse (1987), "The CES is a discrete choice model?", *Economics Letters*, 24(2), pp. 139–140.

Anderson, Simon, André de Palma and Jacques Thisse (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press.

Armstrong, Mark (2013), "A more general theory of commodity bundling", *Journal of Economic Theory*, 148(2), pp. 448-472.

Gentzkow, Matthew (2007), "Valuing new goods in a model with complementarity: online newspapers", *American Economic Review*, 97(3), pp. 713–741.

Hanemann, Michael (1984), "Discrete/continuous models of consumer demand", *Econometrica*, 52(3), pp. 541–561.

Hotelling, Harold (1932), "Edgeworth's taxation paradox and the nature of demand and supply functions", *Journal of Political Economy*, 40(5), pp. 577–616.

Jaffe, Sonia and Glen Weyl (2010), "Linear demand systems are inconsistent with discrete choice", *The B.E. Journal of Theoretical Economics*, 10(1) (Advances), article 52.

Jaffe, Sonia and Scott Duke Kominers (2012), "Discrete choice cannot generate demand that is additively separable in own price", *Economics Letters*, 116(11), pp. 129-132.

McFadden, Daniel (1980), "Econometric models for probabilistic choice among products", *Journal of Business*, 53(3), pp. S13-S29.