Identifying combinatorial valuations from aggregate demand

Itai Sher a,*, Kyoo il Kim b,c,*

a University of Minnesota, United States
b Michigan State University, United States
c Sungkyunkwan University, South Korea

Received 14 June 2013; final version received 3 July 2014; accepted 7 July 2014
Available online 14 July 2014

Abstract

We study identification of combinatorial valuations from aggregate demand. Each utility function takes as arguments subsets or, alternatively, quantities of the multiple goods. We exploit mathematical insights from auction theory to generically identify the distribution of utility functions. In our setting, aggregate demand for each item is observable while demand for bundles is not. Nevertheless, our identification result allows us to recover the latter.

© 2014 Elsevier Inc. All rights reserved.

JEL classification: C14; D11; D44

Keywords: Combinatorial valuations; Identification; Distribution of utility functions; Submodularity; Gross substitutes; Combinatorial auctions

* We are grateful to Patrick Bajari and Jeremy Fox for comments and conversations relating to the paper. We would also like to thank the organizers and participants at the conference on “Bounded Rationality, Revealed Preferences and Consumer Demand” at UCL, the 2012 North American Summer Meetings of the Econometric Society, and the Applied Micro Seminar at the University of Minnesota.
* Corresponding authors.
E-mail addresses: isher@umn.edu (I. Sher), kyookim@msu.edu (K. Kim).
1. Introduction

This paper studies identification of the distribution of combinatorial valuations from aggregate demand. The paper is relevant to the concerns of several literatures. First our exercise is in the tradition of the classic literature on the integrability of demand (Antonelli [1], Samuelson [29], Houthakker [14], Hurwicz and Uzawa [15], Mas-Colell [21]). This literature seeks to uncover a single consumer’s utility function from her demand correspondence. However, our analysis occurs in the context of a more recent literature on Walrasian equilibrium with discrete goods and combinatorial auctions (Kelso and Crawford [16], Gul and Stacchetti [12], Lehmann, Lehmann, and Nisan [18], Cramton, Shoham, and Steinberg [6]). This literature deals with complex resource allocation problems involving bundles of discrete goods. By exploring the link between preferences and demand, our results contribute to an understanding of consumer or bidder behavior in such settings. Accordingly, our work departs from the classic integrability literature by identifying valuations from demand when goods are discrete. A further departure is allowed by the discreteness of our setting: Rather than identifying a single utility function from individual demand, we are able to identify the distribution of utility functions in a finite population from aggregate demand. In this respect our exercise is thematically related to the invertibility results for continuous demand systems (e.g. Brown and Matzkin [3,4], Lewbel [19], Beckert and Blundell [2]) that concern recovering distributions of stochastic demand functions or of unobserved consumer heterogeneity. Yet our paper deals with a very different environment – that of discrete goods and combinatorial preferences – and hence necessitates methods and arguments that are very different from those employed in the literature on continuous demand systems. In particular we employ various substitutes notions particularly adapted to the combinatorial setting.

Our results ultimately have relevance to the empirical literature on discrete choice models of demand. A common simplifying assumption in this literature is that a consumer is allowed to purchase only one alternative in a given period. There has been some work, of both a theoretical and an empirical nature, on discrete choices with multiple purchases, going back to Manski and Sherman [20] and Train, McFadden, and BenAkiva [31]. See also Hendel [13], Dube [8], and Gentzkow [11] for more recent works. These studies have relied on parametric functional forms for both distributions of heterogeneity and the class of underlying utility functions. This literature has restricted attention to data schemes where demand for bundles are available. In contrast, our approach studies identification in the situation where 1) individual goods are priced separately, so there are no bundle-specific prices, and 2) only demand for each product is available, not demand for bundles. However, our analysis is conducted in a very different setting than those in the empirical discrete choice literature. While we do not expect our results to transfer directly to the models in that literature, we believe that our arguments and the conditions we identify and study would be relevant to partial identification in such settings. Further discussion of this point can be found in Section 6.2.

We now describe our environment and results a bit more precisely. We assume that there is a finite number of goods and a finite number of consumers. (We can also handle the case of finitely many consumer types, as discussed in the Online Appendix.) Each consumer has a utility for every package of goods. Utility is assumed to be quasi-linear, so that there is a money good that enters linearly into the utility function. At every price vector, we observe aggregate demand. That is to say, we only observe the total number of units of each good demanded; we do not observe how many consumers demand any given package. For example, if one unit of good 1 and one unit of good 2 are demanded, we do not observe whether these two units were demanded by two separate consumers or by only one consumer. Using only the demands of each good, generically
we are able to identify the distribution of utility functions in the population. The precise notions of genericity that we require as well as their rationale can be found in Sections 2.2.2 and 5.1. Our notion of generic identification is related to that of e.g., McManus [22] and Chiappori and Ekeland [5], but we apply the notion in a very different setting. In particular our application in the context of multi-unit demand is novel (see Definition 2) due to the complex structure of the set of valuations there considered.

It is notable that we can identify such a complex object as the distribution of utility functions using only linear prices. Our identification of the distribution of utility functions uses data on only aggregate demand for each product. We can then derive demand for bundles – that is, the joint distribution of product consumptions – from the marginal demands for individual products. Another consequence of our results is that, for example, the goods in question are iPads and iPhones, we can identify the differential welfare effect of a rise in iPad prices on consumers who purchase iPhones, and consumers who do not (at current prices). A third consequence is that from observing demand generated by heterogenous consumers at linear prices, we can predict demand at non-linear prices. These consequences show that our analysis allows us to recover non-obvious but important properties of preferences and demand.

A further point of interest is that our identification results are constructive in the sense that we describe how the distribution of utility functions can be recovered from demand given sufficient price variation. Starting at low prices, and observing substitution patterns as prices rise, we gradually recover the distribution of consumer utility functions.

Our identification arguments depend on the assumption that goods are substitutes for consumers. We establish two identification results that employ different assumptions about the way in which goods are substitutes. Theorem 1 assumes that goods are substitutes in the sense that valuations are submodular, meaning that the marginal value of a good is decreasing in consumption of other goods. Theorem 2 assumes that valuations are $M^2$-concave (read “$M$-natural concave”), a property that is closely related to the extensively studied gross substitutes condition, a comparative statics condition on demand which says that demand for any good is (weakly) increasing in the price of other goods. Submodularity is significantly weaker than $M^2$-concavity. On the other hand, in order to obtain an identification result with submodularity, we require the assumption that each consumer demands at most one unit of each good, although the consumer may simultaneously demand many different goods. In contrast, $M^2$-concavity allows us to derive an identification result with multi-unit demand for each good.

An outline of the paper is as follows. Section 2 introduces the framework and our notion of identification. Section 3 presents the assumptions on valuations required for our first identification result. Section 4 presents this first identification result employing the assumption of submodularity. A detailed example shows the constructive method of obtaining identification. Section 5 presents the second identification result employing $M^2$-concavity and allowing multi-unit demand for each good. Section 6 discusses the relation of our paper to McManus [22] and to the discrete choice literature, as well as extensions of our results. Appendix A presents a result sufficient to prove Theorem 2, which is proven in an Online Appendix. The Online Appendix also discusses extensions and contains additional proofs.

---

1. See in particular, Definitions 1 and 2.
2. The property of $M^2$ concavity is due to Murota and Shioura [26]. The close relation between $M^2$-concavity and the gross substitutes condition is explored by Fujishige and Yang [10], Danilov, Koshevoy, and Lang [7], and Murota and Tamura [28]. In the case of multi-unit demand, $M^2$-concavity is equivalent to the strong substitutes condition of Milgrom and Strulovici [23] (see Section 5).
2. Framework

This section presents both our underlying model of consumer preferences and consumer demand (Section 2.1), and our notion of identification (Section 2.2).

2.1. Consumers, valuations, and bundles

2.1.1. Consumers and bundles

There is a finite collection \( I = \{1, \ldots, m\} \) of consumers and a finite collection \( J = \{1, \ldots, n\} \) of goods. Goods are discrete. \( \mathbb{Z}_n^m \) is the set of \( n \)-tuples of nonnegative integers. For any \( z \in \mathbb{Z}_n^m \), and \( j \in J \), \( z_j \) is the \( j \)-component of \( z \). An element \( z \in \mathbb{Z}_n^m \) is a bundle, and \( z_j \) is the number of units of good \( j \) contained in \( z \). \( 0 = (0, \ldots, 0) \in \mathbb{Z}_n^m \) is the zero bundle containing zero units of each good. \( 1 = (1, \ldots, 1) \) is the bundle containing one unit of each good. For any \( j \in J \), let \( e^j \) be the \( j \)th unit vector, that is, the vector whose \( j \)th component is 1 and all of whose other components are 0. For \( z, z' \in \mathbb{Z}_n^m \), \( z \leq z' \) means that each component of \( z \) is no larger than the same component of \( z' \). So \( z' \) contains at least as many units of each good as \( z \). \( z < z' \) means that \( z \leq z' \) and \( z \neq z' \). So \( z < z' \) requires that \( z' \) contains a strictly larger quantity of some good than \( z \), but allows that some components of \( z \) and \( z' \) are the same.

For each \( j \in J \), there is a finite number \( N_j \geq 1 \) such that no consumer values additional units of good \( j \) once she consumes \( N_j \) units, regardless her level of consumption of other goods. \( N_j \) is consumers’ saturation point for good \( j \). So consumers choose bundles from the set:

\[
\mathbb{B} = \{ z = (z_1, \ldots, z_j, \ldots, z_n) \in \mathbb{Z}_n^m : z_j \leq N_j, \forall j \in J \}
\]

We also define the satiating bundle \( \bar{z} = (\bar{z}_1, \ldots, \bar{z}_n) \) to be the bundle satisfying:

\[
\bar{z}_j = N_j, \quad \forall j \in J.
\]

It is convenient to introduce the following notation:

\[
J_k(z) := \{ j \in J : z_j = k \}, \quad \text{for } k = 0, 1, \ldots, N_j.
\]

\[
J_{N_j}(z) := \{ j \in J : z_j = N_j \},
\]

\[
J_{N_j-k}(z) := \{ j \in J : 0 \leq z_j \leq N_j - k \}, \quad \text{for } k = 0, 1, \ldots, N_j.
\]

So \( J_k(z) \) is the set of goods \( j \) such that bundle \( z \) contains exactly \( k \) units of \( j \), \( J_{N_j}(z) \) is the set of goods \( j \) that are at the saturation point in bundle \( z \), and \( J_{N_j-k}(z) \) is the set of goods \( j \) that are at least \( k \) units beneath satiation. Moreover, define:

\[
V_0(z) := \{0\} \cup \{ e^j : j \in J \setminus J_{N_j}(z) \},
\]

\[
V_1(z) := \{0\} \cup \{ e^j : j \in J \setminus J_0(z) \}.
\]

\( V_0(z) \) contains the zero vector and the unit vectors corresponding to goods \( j \) beneath their saturation point in \( z \). \( V_1(z) \) contains the zero vector and the unit vectors for goods \( j \) such \( z \) contains at least one unit of \( j \).

2.1.2. Valuations

Each consumer has a valuation \( v_i : \mathbb{B} \to \mathbb{R}_+ \). \( v_i(z) \) is the value of bundle \( z \) to consumer \( i \).

Throughout the paper, we impose the following assumptions on valuations.
Assumption 1.

Normalization \( v_i(0) = 0 \). The utility to consuming nothing is zero.

Monotonicity/Free Disposal \( z \leq z' \Rightarrow v_i(z) \leq v_i(z'), \forall z, z' \in \mathbb{B}. \) More goods are better.

It is convenient to define \( \mathbb{B}_0 := \mathbb{B} \setminus \{0\} \) as the set of bundles in \( \mathbb{B} \) other than the zero bundle. Each valuation \( v_i \) can be represented as a point in the Euclidean space:

\[
\mathcal{E} := \mathbb{R}^{\mathbb{B}_0}
\]

(5)

That is, in order to represent a valuation, we must assign a real number – a value – to each bundle in \( \mathbb{B}_0 \), which is precisely what a point in the Euclidean space \( \mathcal{E} \) does. (We do not have to assign a value to \( 0 \) because this has been normalized to zero.) While every valuation \( v_i \) may be viewed as a point in the Euclidean space \( \mathcal{E} \), not every point \( r \) in \( \mathcal{E} \) corresponds to a valuation. For example, \( r \) may violate monotonicity, assigning a higher number to a bundle \( z' \) than to a bundle \( z \) which contains fewer units of each good than \( z' \). Also, normalization and monotonicity together imply that each nonzero bundle is assigned a nonnegative value. Ultimately, we impose additional assumptions, and refer to the resulting set of valuations as \( \mathcal{V} \) (or \( \hat{\mathcal{V}} \) in Section 5).

Accordingly, we can view \( \mathcal{V} \) as a subset of the Euclidean space \( \mathcal{E} \).

2.1.3. Prices and demand

In addition to the discrete goods \( J \), there is a continuous money good. Consumers have quasi-linear utility with respect to the money good. Let \( p \in \mathbb{R}_+^n \) be a price vector, where \( p_j \) – the \( j \)th component of \( p \) – is the price of good \( j \). As in the case of bundles, for any pair of price vectors, \( p \) and \( p' \), \( p \leq p' \) means that \( p_j \leq p'_j \) for all \( j \in J \), and \( p < p' \) means that \( p \leq p' \) and \( p \neq p' \).

Given quasi-linearity, a consumer’s utility to consuming bundle \( z \in \mathbb{B} \) at price vector \( p \) is:

\[
v_i(z|p) := v_i(z) - \sum_{j=1}^{n} p_j z_j.
\]

(6)

We may use (6) rather than \( v_i(z|p) := v_i(z) - \alpha_i \sum_{j=1}^{n} p_j z_j \) where \( \alpha_i \in \mathbb{R}_{++} \) by a normalization of the utility function. The individual demand function \( D^i(p) \) for consumer \( i \) is:

\[
D^i(p) := \text{argmax}\{ v_i(z|p) : z \in \mathbb{B} \}.
\]

(7)

Because \( \mathbb{B} \) is a finite set, \( D^i(p) \) is never empty. \( D^i(p) \) is said to be multi-valued at prices \( p \) where \( D^i(p) \) contains more than one bundle, and single-valued at prices where \( D^i(p) \) contains only a single bundle. Due to the discreteness of goods, \( D^i(p) \) will inevitably be multi-valued at some price vectors \( p \) at which the consumer becomes indifferent in choosing among several bundles. Aggregate demand \( D(p) \) is the sum of individual demands:

\[
D(p) := \sum_{i \in I} D^i(p),
\]

(8)

\( D(p) \) specifies the total quantity of each good demanded at every price vector \( p \). Given that individual demand may be multi-valued, the summation in (8) is the Minkowski sum: for any two subsets \( A \) and \( B \) of \( \mathbb{Z}_+^n \), \( A + B := \{ a + b : a \in A, b \in B \} \). This coincides with the ordinary sum when \( A \) and \( B \) are singletons (\( A = \{a\} \) and \( B = \{b\} \)). When individual demands are single-valued, so is aggregate demand, but when any individual demand is multi-valued, aggregate demand is also multi-valued. While individual demands are always contained in \( \mathbb{B} \), aggregate demand may be contained in \( \mathbb{Z}_+^n \) but not in \( \mathbb{B} \) because the total number of units of a good demanded by all consumers may exceed the satiation point of each individual consumer.
2.2. Identification

Here we describe our identification exercise. We discuss several related notions of identification. Identification at a valuation profile is a building block for both (i) identification everywhere and (ii) generic identification. Identification everywhere is perhaps the first notion which would come to mind, but is too strong for our purposes (Section 2.2.1). So we ultimately use generic identification (Section 2.2.2). For expositional simplicity, we assume that the number $m$ of agents and the satiation points $N_j$ are fixed and known to the econometrician. The Online Appendix discusses how, when not known a priori, $m$ and $N_j$ can be identified.

2.2.1. Identification everywhere

Let $\mathcal{V}$ be the set of consumer valuations considered possible by the econometrician. $\mathcal{V}$ is characterized by Assumption 1 and either Assumption 2 (to be presented in Section 3.1) or Assumption 3 (in Section 5). Recall that we assume a finite population of consumers in the set $I = \{1, \ldots, m\}$. Each consumer has a valuation $v_i \in \mathcal{V}$. The object of interest for identification is the profile of valuations $(v_1, \ldots, v_m)$ of the $m$ consumers (up to permutations of the identities — that is, indices — of the consumers). We observe aggregate demand $D(p)$ at all price vectors $p \in \mathbb{R}_+^m$. The same population $I$ of consumers with valuation profile $(v_1, \ldots, v_m)$ is held fixed throughout all observations.

We do not observe how aggregate demand is decomposed into demands of individual consumers, nor do we observe the aggregate demands for bundles of goods. For example, suppose $D(5, 4) = (1, 1)$; in other words, at price vector $(5, 4)$, one unit of good 1 is demanded and one unit of good 2 is demanded. There are two essentially distinct possible explanations for this:

(i) Consumer 1 demands one unit of good 1 $[D^1(5, 4) = (1, 0)]$, and consumer 2 demands one unit of good 2 $[D^2(5, 4) = (0, 1)]$. So $D^1(5, 4) + D^2(5, 4) = (1, 0) + (0, 1) = (1, 1) = D(5, 4)$.

(ii) Consumer 1 demands one unit of each good $[D^1(5, 4) = (1, 1)]$, and consumer 2 demands nothing $[D^2(5, 4) = (0, 0)]$. Again: $D^1(5, 4) + D^2(5, 4) = (1, 1) + (0, 0) = (1, 1) = D(5, 4)$.

We do not directly observe which scenario — (i) or (ii) — generates the observed aggregate demand.

By observing aggregate demand $D(p)$ at all price vectors $p$, we would like to identify the underlying profile of valuations $(v_1, \ldots, v_m)$ generating the demand. Let $\mathcal{V}^m$ be the $m$-fold product of $\mathcal{V}$. For any valuation profile $v = (v_1, \ldots, v_m) \in \mathcal{V}^m$, let $D_v$ be the aggregate demand correspondence generated by $v$ via (7)–(8). Say that valuations are identified from demand at $v$, if given the observed demands $(D_v(p) : p \in \mathbb{R}_+^m, v = (v_1, \ldots, v_m)$ is the unique (up to permutations of consumer indices $i$) profile of valuations in $\mathcal{V}^m$ solving (7)–(8) for all price vectors $p$, where in (7)–(8), $D_v(p)$ plays the role of $D(p)$. Say that valuations are everywhere identified from demand if valuations are identified from demand at $v$ for all $v \in \mathcal{V}^m$.

Suppose, however, that for some valuation profiles $v \in \mathcal{V}^m$, valuations are identified from demand at $v$, and for other valuation profiles $v' \in \mathcal{V}^m$, valuations are not identified from demand at $v'$. Then valuations are not identified everywhere from demand. This will indeed be the case under our specific assumptions as will be shown in Example 1. However, ideally, we would like to quantify the relative importance of those cases where identification succeeds and those where it fails. This is what the notion of generic identification allows us to do.
2.2.2. Generic identification

This section introduces and motivates the notion of generic identification used in our first main result, Theorem 1. This notion is related to that of e.g. McManus [22] and Chiappori and Ekeland [5], but applied in a different setting (see Section 6.1 for a discussion). Our other main result, Theorem 2, uses a variation on this notion (see Section 5). For the remainder of this section, we mean by \( \mathcal{V} \) the set of valuations assumed for Theorem 1 (defined below in Section 3.1). As explained in Section 2.1.2, any valuation \( v_j \) can be represented as a point in the Euclidean space \( E \) defined by (5). So the set of valuations \( \mathcal{V} \) is a subset of \( E \). As will be shown in Section 3.3, the set of valuations \( \mathcal{V} \) assumed for Theorem 1 will have infinite Lebesgue measure. So the product \( \mathcal{V}^m \) is an infinite Lebesgue measure subset of the higher dimensional Euclidean space \( E^m \).

**Definition 1.** Let \( \mathcal{X} \) be the set of all valuation profiles \( v \in \mathcal{V}^m \) such that valuations are not identified from demand at \( v \) (see Section 2.2.1). Then valuations are generically identified from demand if \( \mathcal{X} \) has Lebesgue measure zero (in \( E^m \)).

One interpretation of generic identification is as follows: Suppose the \( m \) valuations comprising the finite population which is the object of identification are selected iid according to some probability measure \( \mu \) with support \( \mathcal{V} \) such that \( \mu \) is absolutely continuous with respect to Lebesgue measure on \( E \). Then, under generic identification, the probability of selecting a valuation profile for which identification fails is zero. In this thought-experiment, once the population is drawn from \( \mu \), the population is held fixed as prices vary as in Section 2.2.1. The assumption that the valuations of different members of the population are chosen iid is actually irrelevant to the thought experiment. If the valuation profile is chosen according to a probability measure \( v \) with support \( \mathcal{V}^m \) such that \( v \) is absolutely continuous with respect to Lebesgue measure on the higher-dimensional Euclidean space \( E^m \), then under generic identification, the probability that identification will fail is zero.

3. Assumptions on valuations: submodularity and one unit per good

3.1. Assumptions

For identification, we impose assumptions on the set of valuations \( \mathcal{V} \) in addition to those in Assumption 1. Section 5 attains another identification result under alternative assumptions.

**Assumption 2.**

**One Unit Per Good (1UPG)** \( N_j = 1, \forall j \in J \). No consumer cares for more than one unit of any good; however, a consumer may consume one unit each of many different goods simultaneously.

**Submodularity** \( v_i(z + e^j + e^\ell) - v_i(z + e^\ell) \leq v_i(z + e^j) - v_i(z), \forall z \in \mathbb{B}, \forall j, \ell \in J \setminus J_e(z) \) with \( j \neq \ell \). The marginal value of consuming a unit of any good is decreasing in the consumption of other goods.

We now discuss the assumptions, starting with 1UPG. 1UPG implies that:

\[ \mathbb{B} = \{0, 1\}^n, \] (9)
since each consumer always desires one or zero units of each good. However, 1UPG allows for complex interactions among different goods in consumption. This contrasts with many models in the discrete choice literature which impose unit demand preferences, and thereby rule out such interactions. Section 5 attains an identification result without the 1UPG assumption.

Our most important assumption is submodularity, a substitutes assumption imposing a kind of diminishing marginal utility. However, submodularity is consistent with certain kinds of complementarity, as will be briefly discussed in Section 3.2. Submodularity says that the incremental value of good $j$ diminishes when a unit of another good $\ell$ is added to the bundle. If $z$ and $z'$ are two bundles in $B$ such that $z \leq z'$ (i.e., $z'$ contains at least as many units of each good as $z$) and $z_j = z'_j$, iterated applications of submodularity imply that the incremental value of a unit of good $j$ to package $z$ is at least as large as its marginal value to package $z'$. In the presence of 1UPG, the interpretation of submodularity can be strengthened: It says that the marginal value of any good is decreasing in the set of other goods consumed. This is because under 1UPG, a consumer will never value more than one unit of any good. Submodular valuations feature prominently in the literature on combinatorial auctions (Lehmann, Lehmann, and Nisan [18]), although submodularity has older roots and features prominently in the literature on combinatorial optimization. (See Fujishige [9] for a thorough treatment.)

3.2. Examples of submodular valuations

This section gives examples of submodular valuations assuming 1UPG so that (9) holds.

Unit demand valuations are valuations in which each good is assigned a value, and a consumer evaluates a package on the basis only of the most desirable good in the package, so that a consumer would never demand more than one good at positive prices. This is the most common assumption in applied discrete choice models. Formally, a valuation $v_i$ on $\{0, 1\}^n$ is a unit demand valuation if there exist nonnegative numbers $\alpha_{ij}$ for $j = 1, \ldots, n$ such that: $v_i(z) = \max\{\alpha_{ij}z_j : j = 1, \ldots, n\}$ for all $z \in \{0, 1\}^n$. At any price vector $p$, a unit demand valuation induces the consumer to select a single item $j$ maximizing $\alpha_{ij} - p_j$, unless $\alpha_{ij} - p_j$ is negative for all $j$, in which case the consumer demands nothing.

Additive valuations are valuations in which the value of a package is the sum of the values of the goods in the package. A valuation $v_i$ on $\{0, 1\}^n$ is additive if there exist nonnegative numbers $\alpha_{ij}$ for $j = 1, \ldots, n$ such that $v_i(z) = \sum_{j=1}^n \alpha_{ij}z_j$ for all $z \in \{0, 1\}^n$. So, the consumer’s valuation for different goods is independent of one another. Therefore at a price vector $p$, the consumer purchases good $j$ whenever: $\alpha_{ij} \geq p_j$, and may purchase many goods.

Any concave transformation of an additive valuation is submodular. Let $u : \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing concave function and let $\alpha_{ij}$ be a nonnegative constant for $j = 1, \ldots, n$. A valuation $v_i$ such that $v_i(z) = u(\sum_{j=1}^n \alpha_{ij}z_j)$ is submodular. If $u$ is strictly concave, the incremental value of any good is strictly decreasing in the set of other goods consumed.

The gross substitutes condition for a valuation is defined in terms of a comparative statics condition on the induced demand correspondence: If the prices of some set of goods go up, and the prices of all other goods remain the same, then demand for goods whose prices have stayed the same does not go down. Formally this condition is written as: $\forall p, p' \in \mathbb{R}_+^n, p \leq p' \Rightarrow \forall z \in D^i(p), \exists z' \in D^j(p'), \forall j \in J, p_j = p_j' \Rightarrow z_j \leq z_j'$. Unit demand valuations, additive valuations, and concave transformations of additive valuations are gross substitutes valuations. Gul and Stacchetti [12] showed that all gross substitutes valuations are submodular. As discussed in Section 3.3, the class of gross substitutes valuations is much narrower than that of submodular valuations. This means in particular that submodular valuations allow for some
forms of complementarity in the sense that submodular valuations may violate the gross substitutes condition: With more than two goods, it is possible, under a submodular valuation, that a rise in the price of good \( j \) (with other prices fixed) may reduce demand for good \( k \) with \( j \neq k \). This, however, must be because of the intermediate substitution of a good \( \ell \) for \( j \) which is a also substitute for \( k \). Sher [30] provides a characterization of submodular valuations in terms of a comparative statics of demand condition that allows such intermediate substitutions.

3.3. Non-degeneracy

This section discusses an existing result that establishes that the set of submodular valuations constitute a non-degenerate subset of the set of all possible valuations for discrete goods satisfying 1UPG. This is critical for our main identification result Theorem 1. As explained in Section 2.1.2, any valuation can be viewed as a point in the Euclidean space \( E \) defined by (5).

**Proposition 1.** (Lehmann, Lehmann, and Nisan [18].) Assume 1UPG, so that \( B \) satisfies (9). Let \( \mathcal{V} \) be the set of valuations satisfying normalization, monotonicity, and submodularity. Then \( \mathcal{V} \) has positive (indeed, infinite)\(^3\) Lebesgue measure in \( E \).

This result establishes that there is a broad class of valuations satisfying submodularity. In contrast, Lehmann, Lehmann, and Nisan [18] also showed that the set of gross substitutes valuations has zero Lebesgue measure in \( E \), so there is a sense in which the set of submodular valuations is much larger than the set of gross substitutes valuations. This latter result on gross substitutes will become relevant for our second identification result, Theorem 2, in Section 5.

4. Identification with submodular valuations and single unit demand for multiple goods

4.1. Overview

This section presents and proves our main identification result:

**Theorem 1.** Under Assumptions 1 and 2, valuations are generically identified from demand.

Identification comes from prices where some consumer becomes indifferent among several bundles. These are prices where demand \( D(p) \) becomes multi-valued and around which there is a discrete jump in demand. The proof will have to overcome several challenges:

**Problem 1** When demand becomes multi-valued, how do we know whether one or many consumers are indifferent?

**Problem 2** If demand \( D(p) \) is multi-valued due to the indifference of one consumer, \( p \) gives the marginal values of some items for some consumer at some bundle \( z \). How do we know which bundle \( z \)?

**Problem 3** If we know (a) marginal values for consumer \( i \) at bundle \( z \) from \( D(p) \), and (b) marginal values for consumer \( i' \) at bundle \( z' \) from \( D(p') \), how do we know whether \( i = i' \)??

\(^3\) Lehmann, Lehmann, and Nisan [18] show that \( \mathcal{V} \) has positive measure, but because \( \mathcal{V} \) is a cone, it follows that it must have infinite measure.
We now outline the proof, explaining the contribution of the solution of each of the above problems. Section 4.4 illustrates the solution with example. Section 4.5 presents the proof.

To identify the distribution of valuations, for each consumer i, we must identify the values $v_i = (v_i(z) : z \in \{0, 1\}^n)$. For bundle z, $n_z$ is the total number of units of all goods contained in z. We can associate z with the set $\{j_i^1, \ldots, j_i^k, \ldots, j_i^n\}$ of goods contained in z. For $k = 1, \ldots, n_z$, let $z^k$ be the bundle that contains only the first k goods in z, that is, the goods, $\{j_i^1, j_i^2, \ldots, j_i^k\}$, so that, in particular, $z^0 := 0$. We may express $v_i(z)$ as a telescoping sum:

$$v_i(z) = v_i(z^0) + \sum_{k=0}^{n_z-1} [v_i(z^{k+1}) - v_i(z^k)] = \sum_{k=0}^{n_z-1} [v_i(z^k + e_{j_i}^k) - v_i(z^k)], \quad (10)$$

where we have used the normalization $v_i(0) = 0$. So to recover $v_i$, it is sufficient to recover all terms of the form $v_i(z^0 + e_{j_i}^k) - v_i(z^0)$, where $z^0 \in \{0, 1\}^n$ and $j_i \in J_0(z^0)$, where, recall that $J_0(z^0)$ is the set of goods not in $z^0$. In other words, to recover i’s valuation, it is sufficient to recover i’s marginal value of adding to any package $z^0$ any good $j$ not already contained in it.

For any bundle $z'$ and $j \in J_0(z')$, assume for the moment that there exists a price vector $p(z', j)$ satisfying $\{z', z' + e_j\} \subseteq D(p(z', j))$. Then $p(z', j)$ makes $z'$ optimal for $i$, and, moreover, makes $i$ indifferent between $z'$ and $z' + e_j$, implying that:

$$p(z', j) = v_i(z' + e_j) - v_i(z'), \quad (11)$$

where $p(z', j)$ is the j-component of $p(z', j)$, giving the price of good $j$. $p(z', j)$ identifies $i$’s marginal value for good $j$ at package $z'$ via (11). The collection $(p(z', j) : z' \in \{0, 1\}^n, j \in J_0(z'))$ of all such price vectors for consumer $i$ then identifies $i$’s valuation $v_i$ via (10).

However, we do not directly observe individual demand, but only aggregate demand, which combines the demands of many consumers. So even if we have a price vector $p(z', j)$ satisfying (11) in hand, we will only observe that $\{y, y + e_j\} \subseteq D(p(z', j))$, where $y \in \mathbb{Z}_+^n$ is a bundle that combines the demand of consumer $i$ with the demands of all other consumers. This is why we will require a solution to Problem 2 to separate out within $y$, $i$’s demand for $z'$ from the demands of all other consumers. However, even if we solve this problem, we will only attain a single marginal value of the form (11), whereas to recover $i$’s entire valuation we require many such marginal values, and hence a solution to Problem 3.

To solve Problems 2 and 3, our proof strategy actually is to search for something stronger than price vectors of the form $p(z', j)$: We acquire price vectors $p(z', +)$ satisfying

$$D_+^l(p(z', +)) = z' + V_0(z'), \quad (12)$$

where we have used the Minkowski sum defined in Section 2.1. Recall from (3), that $V_0(z')$ is the set containing the zero vector and the unit vectors corresponding to all goods beneath their satiation point in $z'$. Because we assume 1UPG, this means that $V_0(z')$ contains exactly the zero vector and the unit vectors corresponding to all goods not contained in $z'$. If $p(z', +)$ induces $i$’s demand to take the form (12), then not only does $p(z', +)$ make $z'$ optimal for $i$, but $p(z', +)$ simultaneously makes $i$ indifferent to adding any good not contained in $z'$, so that: $p_j(z', +) = v_i(z' + e_j) - v_i(z')$, $\forall j \in J_0(z')$. As discussed above, we do not observe individual demand, but only aggregate demand, so that we must disentangle the individual demand (12) from the aggregate demand that we observe. Moreover, because individual demand of the form (12) will typically involve indifference among many bundles (more than just two bundles), this multiplicity will be passed over to aggregate demand, so that separating individual demand from
aggregate demand will involve answering the question of whether one or many consumers are indifferent at price vector \( p(z', +) \), or, in other words, a solution to Problem 1 (in addition to Problems 2 and 3). Moreover we must establish:

Prerequisite 0. Price vectors \( p(z', +) \) satisfying (12) exist.

So if we establish Prerequisite 0, and solve Problems 1–3, thereby finding the collection of price \( P_+ = (p(z, +) : z \in \{0, 1\}^n) \), we will have established Theorem 1.

4.2. Sufficiency of observing single-valued demand only

While the above argument seems to depend on observing aggregate demand where it becomes multi-valued, we do not literally need to observe demand where it becomes multi-valued to achieve identification. Indeed, if we observed demand everywhere where it was single valued, we could still obtain identification. If demand is multi-valued at price vector \( p \), then this can be viewed merely as a shorthand for a certain pattern of substitution as prices vary in a neighborhood around \( p \). Arbitrarily small price changes can break the indifference in favor of any of the bundles demanded at \( p \), so that by looking at price vectors near \( p \) where demand is single-valued, we can recover the multi-valued demand at \( p \). The Online Appendix makes this precise and shows that it is sufficient for our results.

4.3. Role of genericity

Theorem 1 establishes generic identification rather identification everywhere. The following example illustrates why identification everywhere within \( Y_m \) is impossible.

Example 1. Assume two goods \( (J = \{1, 2\}) \) and two consumers Ann and Bob, Scenarios 1 and 2 cannot be told apart as they generate the same aggregate demand at all prices:

\[
\begin{align*}
\text{Scenario 1} & \quad \text{Ann has the additive valuation} \quad v_a(1) = 1, \quad v_b(2) = 2, \quad v_a([1, 2]) = v_a(1) + v_a(2) = 3. \\
& \quad \text{Bob has the additive valuation} \quad v_b(1) = 5, \quad v_b(2) = 3, \quad v_b([1, 2]) = v_b(1) + v_b(2) = 8. \\
\text{Scenario 2} & \quad \text{Ann has the additive valuation} \quad v_a(1) = 1, \quad v_a(2) = 3, \quad v_a([1, 2]) = v_a(1) + v_a(2) = 4. \\
& \quad \text{Bob has the additive valuation} \quad v_b(1) = 5, \quad v_b(2) = 2, \quad v_b([1, 2]) = v_b(1) + v_b(2) = 7.
\end{align*}
\]

While Example 1 rules out identification everywhere, it does not rule out generic identification. We show that the class of valuation profiles disallowing identification has measure zero. Unlike the valuations in Example 1, almost every valuation in \( \mathcal{V} \) satisfies:

\textbf{Strict Submodularity.}

\[
v_j(z + e^j + e^\ell) - v_j(z + e^\ell) < v_j(z + e^j) - v_j(z),
\forall z \in \mathbb{B}, \forall j, \ell \in J \setminus J_v(z) \text{ with } j \neq \ell.
\]

The marginal value of consuming a unit of any good is strictly decreasing in the consumption of other goods.

The additive valuations in Example 1 violate strict submodularity: For additive valuations, the incremental value of a good is independent of consumption of other goods. Note however, that
there exist strictly submodular valuations that are arbitrarily close to additive valuations (using Euclidean distance in $\mathcal{V}$). For example, one can make a concave transformation of an additive valuation $v_i(z) = u (\sum_{j=1}^{n} \alpha_{ij} z_j)$ arbitrarily close to additive by choosing the concave function $u$ to be arbitrarily close to linear on the interval $[0, \sum_{j=1}^{n} \alpha_{ij}]$. Our generic identification result uses the fact that valuations are almost surely strictly submodular. This implies that for no consumer are the values of any two goods completely independent.

Genericity also allows us to impose assumptions on valuation profiles. For example, given any probability measure $\mu$ with support $\mathcal{V}$ such that $\mu$ is absolutely continuous with respect to Lebesgue measure on $\mathcal{E}$ (see Section 2.2.2), if $m$ valuations are selected i.i.d. from $\mu$:

- With probability 1, if $i$ is indifferent about buying $j$ at prices $p$ (given $i$’s other optimal purchases), no other consumer is indifferent about buying $j$ at $p$. \hfill (14)

However, this by itself does not solve Problem 1, for example, because $\mu$ may be such that:

- With positive probability, there exists a price vector $p$ such that:
  1. Consumer $i$ is indifferent about purchasing good $j$, and
  2. Consumer $k$ is indifferent about purchasing good $\ell$, where $i \neq k$ and $j \neq \ell$. \hfill (15)

So when demand is multi-valued demand, we require a stronger argument to determine whether one or many consumers are indifferent.

4.4. Example illustrating Theorem 1

Here we illustrate the argument for Theorem 1 with a simple example. Assume two goods, 1 and 2, and a finite number $m$ of consumers. Ann is one such consumer, with valuation:

<table>
<thead>
<tr>
<th>Bundle</th>
<th>(1, 0)</th>
<th>(0, 1)</th>
<th>(1, 1)</th>
</tr>
</thead>
</table>
| Ann    | 3      | 5      | 7      | \hfill (16)

Let $v_a$ be Ann’s valuation. $v_a(0, 0) = 0$. Ann does not value a second unit of either good. Ann’s valuation satisfies Assumption 2. In particular, it is submodular, as good 1 is less valuable once good 2 is possessed (and vice versa). Indeed, Ann’s valuation is strictly submodular on $[0, 1]^n$, a property holding generically relative to the submodular valuations, as discussed above.

Fig. 1 is an alternative representation of Ann’s valuation. Each edge is labeled with a marginal value, corresponding to the difference in Ann’s value between the two bundles the edge connects. So the edge connecting $(1, 1)$ and $(1, 0)$ gives Ann’s marginal value for good 2, when she possesses good 1. The edge is labeled 4 because $v_a(1, 1) - v_a(1, 0) = 7 - 3 = 4$. Strict submodularity
is represented by the fact that for every pair of parallel edges, the bottom edge is labeled by a larger number than the top edge. This is because parallel edges represent addition of the same good, and an edge higher in the diagram represents addition of the good to a larger package. Fig. 1 represents Ann’s valuation via marginal values as in (10).

Let \( z = (z_1, z_2) \) be a bundle in Fig. 1. We define a price vector \( p(z, +) \). If \( z_1 = 0 \), the first component of \( p(z, +) \), \( p(z, +)_1 \), is the number labeling the edge connecting \( z \) to \( z + (1, 0) \). If \( z_1 = 1 \), let \( p(z, +)_1 = 0 \). \( p(z, +)_2 \) is defined similarly. If \( z = (0, 0) \), \( p(z, +)_1 = 3 \) and \( p(z, +)_2 = 5 \). \( p(z, +) \) is the price vector whose \( j \)th component is equal to Ann’s marginal value for adding a unit of the \( j \)th good to package \( z \). \( p(z, +) \) will turn out to be a price vector with the property given in (12) of Section 4.1, with \( i = a \) (\( a \) being the index for Ann). The notation “+” hints that \( p(z, +) \) provides Ann’s values for adding goods to \( z \). We also need a price vector \( p(z, -) \) encoding Ann’s values for subtracting goods from \( z \). If \( z_1 = 1 \), \( p(z, -)_1 \) is the number labeling the edge connecting \( z \) and \( z - (1, 0) \). If \( z_1 = 0 \), then let \( p(z, -)_1 = M \) be a very large number, sufficiently large that when the price of good 1 is \( M \), demand for good 1 is always zero. Effectively, we can think of \( M \) being infinity, corresponding to the infeasibility of subtracting a unit of good 1 from a package containing zero units of good 1. \( p(z, -)_2 \) is defined similarly. For example, if \( z = (1, 1) \), then \( p(z, -)_1 = 2 \) and \( p(z, -)_2 = 4 \).

When \( z = (0, 0) \), \( p(z, +) = (3, 5) \), generating an individual demand for Ann of:

\[
D^a(p((0,0),+)) = D^a(3,5) = \{(0,0),(1,0),(0,1)\} = (0,0) + V_0(0,0).
\]  

So the zero bundle is optimal for Ann, and Ann is indifferent to adding either good (but not both goods). For \( z = (0, 0) \), \( p(z, +) \) has the structure given by (12) of Section 4.1. This establishes Prerequisite 0 for \( z = (0, 0) \) and \( i = a \), asserting existence of a price vector satisfying (12). \( p(z, +) \) induces Ann to be indifferent between \( z, z + (1, 0), \) and \( z + (0, 1) \) by construction. But why does \( p(z, +) \) make \( z = (0, 0) \) optimal? This is the nontrivial part of Prerequisite 0. The answer is submodularity. If Ann’s value for \( (1,1) \) had been 9 instead of 7, then the two edges into \( (1,1) \) in Fig. 1 would have been labeled 6 (the edge connecting \( (1,0) \) to \( (1,1) \)) and 4 (that connecting \( (0,1) \) to \( (1,1) \)) instead of 4 and 2. For either pair of parallel edges in Fig. 1, the higher edge would be labeled with a larger number, meaning the incremental value of a good would increase in consumption of other goods, an assumption known as supermodularity. So Ann’s unique optimal bundle at \( p(z, +) = (3, 5) \) would be \( (1, 1) \), but \( p(z, +) \) is the unique price vector making Ann indifferent among the bundles in \( z + V_0(z) \) when \( z = (0, 0) \). So no price vector would induce Ann to demand \( z + V_0(z) \) for \( z = (0, 0) \), violating Prerequisite 0.

We have argued that \( p(z, +) \) induces Ann’s individual demand to have the correct form. Next we consider aggregate demand, which is what we observe. We now invoke genericity: Using (14), when Ann is indifferent to the prospect of adding either good – as she is at price vector \( p(z, +) = (3, 5) \) – no other consumer will be indifferent about adding either good. Hence all other consumers will have single-valued demand. So aggregate demand will take the form

\[
D(3,5) = \{y,y + (1,0),y + (0,1)\} = y + V_0(y)
\]  

for some bundle \( y \). The fact that demand is multi-valued at prices \((3,5)\) reflects the fact that at these prices, Ann is indifferent among several bundles, and as prices vary around \((3,5)\), arbitrarily small changes in price will break the tie in one way or another.

We have just argued that if Ann’s valuation is given by (16), then aggregate demand at price vector \((3,5)\) will be of the form (18). However, if we observe that aggregate demand takes the form (18) at prices \((3,5)\), can we infer that there is only a single consumer who is indifferent among several bundles at \((3,5)\)? After all, (15) reminds us that two distinct consumers may be
indifferent about adding different goods at the same price vector. This is Problem 1 of Section 4.1. Moreover, can we infer that this single consumer is indifferent exactly among the three bundles (0, 0), (1, 0), and (0, 1)? This would require a solution to Problem 2 as well.

Indeed, we can solve these problems. If two consumers were indifferent at prices (3, 5) such that one was indifferent about adding a unit of good 1 (at his optimum bundle) and the other was indifferent about adding a unit of good 2, then aggregate demand would be of the form:

\[ D(3, 5) = \{ y, y + (1, 0), y + (0, 1), y + (1, 1) \}. \]  

(19)

The addition of bundle \( y + (1, 1) \) stems from the possibility that both indifferent consumers will break the tie by consuming an additional good. So (18) reveals that only one consumer is indifferent, solving Problem 1 at prices (3, 5). 1UPG (see Assumption 2) implies that if at positive prices, a consumer is indifferent among three potentially demanded bundles of the form \( z, z + (1, 0), \) and \( z + (0, 1), \) these bundles must be \((0, 0), (1, 0), \) and \((0, 1), \) solving Problem 2. (Note that solving Problem 2 when there are more than two goods is more difficult.)

We chose Ann’s valuation so that her demand took form (17). If we had chosen it so that:

\[ D^a(3, 5) = \{ (0, 0), (1, 0), (0, 1), (1, 1) \}. \]  

(20)

then aggregate demand would take form (19), and we could not have known whether one or two consumers were generating the indifference. But Ann’s demand can only have form (20) if her valuation is additive. Generically, relative to our assumptions, Ann’s valuation will be strictly submodular, and so not additive. Strict submodularity implies that at prices where Ann is indifferent between \((0, 0)\) and \((0, 1)\), she strictly prefers \((1, 0)\) to \((1, 1)\), because consumption of good 1 reduces the marginal value of good 2. So at such prices, \((1, 1)\) cannot be optimal.

To summarize: By observing aggregate demand of form (19), we infer that at prices (3, 5), there is a single consumer \( i \) indifferent among \((0, 0)\), \((1, 0)\), and \((0, 1)\). So we can infer that:

\[
\begin{align*}
  v_i(0, 0) &= 0 \\
  v_i(1, 0) &= \left[ v_i(1, 0) - v_i(0, 0) \right] + v_i(0, 0) = 3 \\
  v_i(0, 1) &= \left[ v_i(0, 1) - v_i(0, 0) \right] + v_i(0, 0) = 5 \\
  v_i(1, 1) &= ?
\end{align*}
\]  

(21)

where we have used our normalization \( v_i(0, 0) = 0 \). In order to completely identify Ann’s valuation, we must still uncover \( v_i(1, 1) \). This will require a solution to Problem 3.

To recover \( v_i(1, 1) \), we appeal to prices \( p(z, -) = (2, 4) \). Recall that \( p(z, -) \) gives Ann’s marginal value for removing either good from \( z = (1, 1) \). At this price vector, Ann’s demand is: \( D^a(p((1, 1), -)) = ((1, 1), (1, 0), (0, 1)) = (1, 1) - V_i(1, 1) \). For submodular valuations, both \( p(z, -) \) and \( p(z, +) \) make bundle \( z \) optimal. While \( p(z, +) \) also makes it optimal to add any good not contained in \( z \), \( p(z, -) \) makes it optimal to remove any good contained in \( z \). As above, we argue that generically at these prices aggregate demand has form:

\[ D(2, 4) = \{ y', y' - (1, 0), y' - (0, 1) \} = y' - V_i(y') \]  

(22)
for some bundle $y'$. As above, we argue that given (22), all indifference is generated by a single consumer $i'$, indifferent among bundles $(1, 1)$, $(1, 0)$, and $(0, 1)$, and so has valuation satisfying:

$$v_i'(1, 1) - v_i'(0, 0) = 4, \quad \text{and} \quad v_i'(1, 1) - v_i'(0, 1) = 2. \quad (23)$$

To uncover $v_i(1, 1)$, and so Ann’s valuation, we must be able to infer that consumer $i$ whose indifference we observe at prices $(3, 5)$, yielding (21) is the same as consumer $i'$ whose indifference we observe at $(2, 4)$, yielding (23). This would amount to a solution to Problem 3.

The key to identification is the following observation: For any consumer $i$,

$$v_i(0, 0) + [v_i(1, 0) - v_i(0, 0)] + [v_i(1, 1) - v_i(1, 0)]$$

$$= v_i(1, 1)$$

$$= v_i(0, 0) + [v_i(1, 1) - v_i(0, 0)] + [v_i(1, 1) - v_i(0, 1)]. \quad (24)$$

Eq. (24) says that if consumer $i$ starts at bundle $(0, 0)$, and $i$ receives one unit of good 1, increasing her utility by $v_i(1, 0) - v_i(0, 0)$, and then receives a unit of good 2, increasing her utility by $v_i(1, 1) - v_i(1, 0)$, her utility will be that of consuming both goods, $v_i(1, 1)$, which is the same as the utility she would have if we had first given her good 2, and then given her good 1. In other words, the order in which the goods are added to the consumer’s bundle does not affect her total utility. Rearranging terms, it is convenient to rewrite (24) as:

$$i = i' \implies [v_i(1, 1) - v_i(0, 0)] - [v_i(1, 0) - v_i(0, 0)]$$

$$= [v_i'(1, 1) - v_i'(1, 0)] - [v_i'(1, 1) - v_i'(0, 1)]. \quad (25)$$

As just argued, the above relationship must hold when $i = i'$. Moreover, it would be a coincidence if this relationship held when $i \neq i'$. Indeed generically we can strengthen (25) to:

$$i = i' \iff [v_i(1, 1) - v_i(0, 0)] - [v_i(1, 0) - v_i(0, 0)]$$

$$= [v_i'(1, 1) - v_i'(1, 0)] - [v_i'(1, 1) - v_i'(0, 1)].$$

In our example, if $i$ is the consumer whose indifference is observed at prices $(3, 5)$ and $i'$ is the consumer whose indifference is observed at prices $(2, 4)$, we have:

$$[v_i'(1, 1) - v_i'(1, 0)] - [v_i'(1, 1) - v_i'(0, 1)]$$

$$= 4 - 2 = 5 - 3$$

$$= [v_i(0, 1) - v_i(0, 0)] - [v_i(1, 0) - v_i(0, 0)]. \quad (26)$$

As we have explained, generically, (26) would not hold if $i$ and $i'$ were distinct consumers. It follows that $i = i'$. This is the solution to Problem 3. Using (21) and (23), we now obtain:

$$v_i(1, 1) = [v_i(1, 1) - v_i(1, 0)] + v_i(1, 0) = [v_i'(1, 1) - v_i'(1, 0)] + v_i(1, 0) = 4 + 3 = 7.$$

In combination with the information in (21), we have now identified the valuation of a single consumer $i$ – who is Ann – in our example. Identification would be completed by applying a similar procedure for all other consumers.

If there are three goods, then rather than a diamond as in Fig. 1, one can represent the consumer’s valuation on a cube. In the diagram of Fig. 2, each vertex corresponds to a consumption bundle, and each edge corresponds to the marginal utility of consuming some good. To avoid cluttering the diagram, we have labeled some – but not all – of the edges. A procedure similar
to the one described above, climbing along the edges of the cube, would allow us to recover the marginal utilities from demand data. For more than three goods, one could construct higher dimensional figures.

4.5. Proof of Theorem 1

This section proves Theorem 1, filling in the details outlined in Section 4.1. The formal argument parallels the development of the example of Section 4.4. Throughout the course of the proof, it will be useful to keep in mind that under 1UPG, $\mathbb{B} = \{0, 1\}^n$, so that for any bundle $z \in \mathbb{B}$, $J_0(z) = J \setminus J_1(z) = J \setminus J_n(z)$, or equivalently, $J_1(z) = J \setminus J_0(z) = J^\prime(z)$. In other words, under 1UPG, any bundle $z$ that may be demanded by some individual contains either one or zero units of each good, and the consumer’s desire for a good $j$ is satiated at package $z$ exactly if the package contains a single unit of the good. These relations will cease to hold when we relax the 1UPG assumption in Section 5.

We now state a result which implies Theorem 1, but which is more amenable to direct proof. As above, $\mathcal{V}$ is the set of all valuation profiles satisfying Assumptions 1 and 2.

**Proposition 2.** Suppose Assumptions 1 and 2 hold. Then there exists a set of valuation profiles $\mathcal{S} \subseteq \mathcal{V}^m$ with the following properties:

1. If $v = (v_1, \ldots, v_i, \ldots, v_{i'}, \ldots, v_n) \in \mathcal{S}$, where $i, i' \in 1$ are such that $i \neq i'$, then $v_i \neq v_{i'}$.
2. $\mathcal{S}$ contains almost all valuation profiles in the sense that $\mathcal{V}^m \setminus \mathcal{S}$ has Lebesgue measure zero in $\mathcal{E}^m$. (Recall that $\mathcal{V}^m$ has infinite Lebesgue measure in $\mathcal{E}^m$.)
3. Choose any valuation profile $v = (v_1, \ldots, v_i, \ldots, v_m) \in \mathcal{S}$. Let $P_+ = \{p(z, +) : z \in \{0, 1\}^n\}$ and $P_- = \{p(z, -) : z \in \{0, 1\}^n\}$ be two collections of price vectors satisfying:

$$p(z, +)_j = 0, \quad \forall j \in J_1(z), \quad (27)$$

$$p(z, -)_j = M, \quad \forall j \in J_0(z), \quad (28)$$

$$p(z + e^j, -)_j = p(z, +)_j, \quad \forall j \in J_0(z), \quad (29)$$

\footnote{See Proposition 1.}
where \( M \) is a number sufficiently large that under valuation profile \( v \), for every good \( j \), aggregate demand for \( j \) is always zero when the price of \( j \) is \( M \). Then the following conditions are equivalent:

(a) For all \( z \in \{0, 1\}^n \),

\[
\exists y(z, +) \in \mathbb{Z}_+^n \text{ such that } D_v(p(z, +)) = y(z, +) + V_0(z), \quad \text{and} \quad (30)
\]

\[
\exists y(z, -) \in \mathbb{Z}_+^n \text{ such that } D_v(p(z, -)) = y(z, -) - V_1(z). \quad (31)
\]

Moreover, for all \( z \in \{0, 1\}^n \),

\[
p(z, +) + p(z + e^j, +) = p(z, +) + p(z + e^{j'}, +), \quad \forall j, j' \in J_0(z) \text{ with } j \neq j'. \quad (32)
\]

(b) There exists a unique consumer \( i \in I \) such that:

\[
p(z, +) = v_i(z + e^j) - v_i(z), \quad \forall z \in \{0, 1\}^n, \forall j \in J_0(z). \quad (33)
\]

The proposition asserts the existence of a set of valuation profiles \( S \), which is generic relative to the set \( \mathcal{V}^m \) of valuation profiles satisfying Assumptions 1 and 2 (condition 2 of the proposition) and which contains only profiles not assigning distinct consumers the same valuation (condition 1). The set \( S \) is precisely defined in Section 4.5.1, Condition 3 – the main clause of the proposition – restricts attention to valuation profiles in \( S \). It asserts the equivalence of two conditions: 3a and 3b. 3a is expressed completely in terms of observables, and can be recovered from the demand correspondence. 3b relates certain critical prices to \( i \)'s marginal utilities of goods to packages and thereby allows one to recover valuations from demand. Both conditions concern two sets, \( P_+ \) and \( P_- \), of price vectors. Such price vectors were discussed more informally in Sections 4.1 and 4.4. Eqs. (30)–(31) assert that these price vectors generate aggregate demand with a certain structure. Eq. (30) asserts that at \( p(z, +) \in P_+ \), aggregate demand becomes multi-valued, including some quantity vector \( y(z, +) \) as well as other quantity vectors, each of which differs from \( y(z, +) \) only by the addition of one unit of some good not contained in \( z \). An instance of this structure was given by (18) for the example of Section 4.4, where \( z = (0, 0) \) and \( p(z, +) = (3, 5) \). Similarly, for \( p(z, -) \in P_- \), (31) asserts that aggregate demand consists of some quantity vector \( y(z, -) \) as well as other quantity vectors each of which differs from \( y(z, -) \) only by the subtraction of one unit of some good contained in \( z \). Eq. (32) asserts an internal relation among the price vectors. (Because of (29), (32) imposes a relation among price vectors in \( P_- \) as well as \( P_+ \).) If one plugs the marginal values from (33) into (32), one can see that (32) is closely related to (24), which was the key to identification in the example, and led to the solution of Problem 3. If one interprets price vectors in \( P_+ \) as representing marginal values of some agent \( i \) as in (33), then (32) says that if one starts with bundle \( z \) and adds two goods not in \( z \), thereby increasing \( i \)'s value by the sum of \( i \)'s marginal values of the each good to the package to which it is added, then in arriving at \( i \)'s value for the new package, the order in which the goods are added does not matter.

We now explain why Proposition 2 implies Theorem 1, so that it is sufficient to prove Proposition 2. Recall that \( S \) is a generic set of valuation profiles by part 2 of Proposition 2. Choose profiles \( v = (v_1, \ldots, v_n) \), \( v' = (v'_1, \ldots, v'_n) \in S \) that are distinct (and remain distinct even if we permute the indices of either profile). To establish Theorem 1, it is sufficient to show that the induced demand correspondences are distinct: \( D_v(\cdot) \neq D_{v'}(\cdot) \). By part 1 of Proposition 2, the unordered sets of valuations \( \{v_1, \ldots, v_n\} \) and \( \{v'_1, \ldots, v'_n\} \) in \( v \) and \( v' \) are distinct and each of
these sets contains \( n \) distinct elements. It follows that there exists \( i \) such that \( v_i \not\in \{v'_i, \ldots, v'_m\} \). Now define the vectors in the set \( P_\delta \) by Eqs. (27) and (33). Similarly define the vectors in \( P_- \) by (28) and \( p(z, -) = v_i(z) - v_i(z - e^j), \forall j \in J(z) \). Monotonicity of \( v_i \) implies that these price vectors have no negative components, and by construction, they satisfy (27)–(29). So we can apply part 3 of Proposition 2. The above implies that \( v_i \) and the price vectors in \( P_+ \) satisfy 3b. So the implication 3b \( \Rightarrow \) 3a implies that \( D_{v_i}(\cdot) \) satisfies 3a. If \( D_{v_i}(\cdot) = D_{v'_i}(\cdot) \), then \( D_{v'_i} \) also satisfies 3a relative to the same set of price vectors \( P_+ \). But then the implication 3a \( \Rightarrow \) 3b implies there exists \( i' \) such that \( v'_{i'} \), like \( v_i \), satisfies (33) with \( P_+ \). But this together with normalization from Assumption 2 implies that \( v'_{i'} = v_i \), a contradiction. So \( D_{v_i}(\cdot) \not= D_{v'_i}(\cdot) \), as we wanted to show. So Proposition 2 implies Theorem 1. Indeed, the proof of Proposition 2 – especially Lemma 6 – shows how to constructively recover individual valuations from aggregate demand.

4.5.1. The generic set \( \mathcal{S} \) of valuation profiles
To define the set \( \mathcal{S} \) from Proposition 2, we proceed in two steps. First we impose conditions on individual valuations \( \mathcal{V} \), thereby assuming that valuations are drawn from the subset \( \mathcal{W} \) of \( \mathcal{V} \) satisfying these conditions (Section 4.5.1.1). Next we impose certain conditions on relations among different valuations and thereby define \( \mathcal{S} \) as a subset of \( \mathcal{W}^m \) (Section 4.5.1.2). The imposed conditions immediately imply that \( \mathcal{S} \) satisfies condition 1 of Proposition 2. Next we prove that \( \mathcal{S} \) satisfies condition 2 of Proposition 2 (Section 4.5.1.3). The interpretation is that both the properties ruled out within individual valuations and the relations ruled out among different valuations correspond to coincidences which are exceedingly unlikely – i.e., have zero probability – if valuation profiles are selected from some continuous distribution on \( \mathcal{V}^m \).

4.5.1.1. Conditions on individual valuations
This section specifies two sets of conditions that define a subset \( \mathcal{W} \) of the valuations \( \mathcal{V} \). The first set of conditions is:

\[
\forall z \in [0, 1]^n, \forall j, j' \in J_0(z) \text{ with } j \not= j'. \\
v_i(z + e^j) - v_i(z) \not= v_i(z + e^j + e^{j'}), \\
v_i(z + e^{j'}) - v_i(z + e^j). \\
(34)
\]

Condition (34) says the incremental value of \( j \) to \( z \) changes once we add another good \( j' \) to \( z \). Combining (34) with submodularity amounts to strict submodularity (13). The next set of conditions is:

\[
v_i(1 - e^i) \not= v_i(1), \quad \forall i \in I, \forall j \in J. \\
(35)
\]

Condition (35) says that removing any good from package 1 affects the agent’s value. Combined with monotonicity from Assumption 1, (35) implies that removing any good from 1 reduces the agent’s value. We now define the promised set \( \mathcal{W} \). \( \mathcal{W} = \{v_i \in \mathcal{V} : v_i \text{ satisfies (34) and (35)}\} \).

Observation 1. \( \mathcal{W} \) is precisely the set of valuations satisfying normalization and 1UPG (Assumption 2), and which, moreover, satisfy strict submodularity (13) and strict monotonicity: \( z < z' \Rightarrow v_i(z) < v_i(z'), \forall z, z' \in [0, 1]^n \).

Strict submodularity was explained above. For strict monotonicity, it is sufficient to show: \( v_i(z) < v_i(z + e^j), \forall z \in [0, 1]^n, \forall j \in J_0(z) \). Choose \( z \in [0, 1]^n \) and \( j \in J_0(z) \). If \( z + e^j = 1 \), the desired inequality follows from (35) and monotonicity (see Assumption 1). If \( z + e^j < 1 \), then, using strict submodularity, \( 0 < v_i(1) - v_i(1 - e^j) < v_i(z + e^j) - v_i(z) \). This establishes \( v_i(z) < v_i(z + e^j) \) and so (13). Conversely, it is straightforward to verify that any valuation satisfying normalization, 1UPG, strict submodularity, and strict monotonicity belongs to \( \mathcal{W} \).
4.5.1.2. Conditions on relations among valuations  We now impose conditions on valuation profiles that, unlike (34)–(35), cannot be reduced to conditions on individual valuations:

\[
 v_{i'}(x) - v_{i'}(y) \neq \sum_{j \in J_0(z)} \left[ v_i(z + e^j) - v_i(z) \right] (x_j - y_j)
\]

\(\forall i, i' \in I, \ with \ i \neq i', \forall x, y, z \in \{0, 1\}^n, \ with \ x \neq y, z \leq x, \ and \ z \leq y. \tag{36}\)

\[
 v_{i'}(x) - v_{i'}(y) \neq \sum_{j \in J_0(z)} \left[ v_i(z) - v_i(z - e^j) \right] (x_j - y_j)
\]

\(\forall i, i' \in I, \ with \ i \neq i', \forall x, y, z \in \{0, 1\}^n, \ with \ x \neq y, x \leq z, \ and \ y \leq z. \tag{37}\)

Condition (36) can be interpreted as saying that when prices are given by \(i\)’s marginal values for adding goods to package \(z\) (assuming that since \(i\) satisfies UPG, the prices for goods already in \(z\) are zero), no other consumer \(i'\) is indifferent among any pair of packages in \(\{0, 1\}^n\) containing \(z\). If consumer \(i'\) were made indifferent among several packages when prices are given by the marginal values of a different consumer \(i\), this would amount to an “unlikely coincidence.” Eq. (37) imposes a similar condition when prices are given for \(i\)’s marginal values from removing goods from package \(z\), (assuming that prices for goods not in \(z\), which then cannot be removed from \(z\) are infinite). Expression (36) gives a set of conditions, each of which asserts a non-equality, the left hand side of which contains an expression which denotes a linear function of \(v_{i'}\) considered as a point in the Euclidean space \(E\), and the right hand side of which contains an expression similarly denoting a linear function of \(v_i\). A similar remark applies to (37). We now define \(S\):

\[
 S = \{ v = (v_1, \ldots, v_m) \in \mathcal{W}^m : v \text{ satisfies (36)--(37).} \} \tag{38}\)

**Proposition 3.** Any valuation profile \(v = (v_1, \ldots, v_m) \in S\) has the following two properties:

\[
 v_i(z + e^j) - v_i(z) \neq v_{i'}(z + e^j) - v_{i'}(z), \quad \forall i, i' \in I \ with \ i \neq i', \forall z \in \{0, 1\}^n, \forall j \in J_0(z). \tag{39}\)

\[
 v_i(z + e^j) - v_i(z + e^{j'}) \neq v_{i'}(z + e^j) - v_{i'}(z + e^{j'}), \quad \forall i, i' \in I \ with \ i \neq i', \forall z \in \{0, 1\}^n, \forall j, j' \in J_0(z) \ with \ j \neq j'. \tag{40}\)

**Proof.** Condition (39) follows from setting \(x = z + e^j\) and \(y = z\) in (36). Condition (40) follows from setting \(x = z + e^j\) and \(y = z + e^{j'}\) in (36). \(\square\)

Condition (39) says that for any package \(z\) and good \(j\) not in \(z\), no two consumers have the same marginal value for good \(j\) at \(z\). Eq. (40) says that for any pair of goods, \(j\) and \(j'\) not in package \(z\), conditional on possessing \(z\), the difference in values of \(j\) and \(j'\) differs across consumers. Condition (39) by itself makes it clear that \(S\) satisfies condition 1 of Proposition 2.

4.5.1.3. Proof that \(S\) satisfies condition 2 of Proposition 2  \(\mathcal{W}\) is derived from \(\mathcal{V}\) via conditions (34)–(35). Conditions (34)–(35) say that valuations \(v_i\), which can be conceived as points in the Euclidean space \(E\) (see Section 3.3) do not lie on certain hyperplanes in \(E\). Any hyperplane in a Euclidean space has zero Lebesgue measure in that space. As \(\mathcal{V} \setminus \mathcal{W}\) is contained in a hyperplane of \(E\), \(\mathcal{V} \setminus \mathcal{W}\) has zero Lebesgue measure in \(E\). Let \(\mathcal{V}_i := \mathcal{V} \setminus \mathcal{W}_i := \mathcal{W}\) for \(i \in I = \{1, \ldots, m\}\). Then \(\mathcal{V}^m \setminus \mathcal{W}^m = \bigcup_{i \in I} (\mathcal{V}_i \times \cdots \times (\mathcal{V}_i \setminus \mathcal{W}_i) \times \cdots \times \mathcal{V}_m)\). So \(\mathcal{V}^m \setminus \mathcal{W}^m\) has zero Lebesgue
measure in the larger Euclidean space $\mathcal{E}^m$. (Proposition 1 implies that both $\mathcal{V}^m$ and $\mathcal{W}^m$ have infinite Lebesgue measure in $\mathcal{E}^m$.) As a valuation can be represented as a point in the Euclidean space $\mathcal{E}$, so a valuation profile can be represented as a point in the Euclidean space $\mathcal{E}^m$. $\mathcal{S}$ is derived from $\mathcal{W}^m$ by (36)–(37), non-equalities which say that valuation profiles $v$ do not lie on certain hyperplanes in $\mathcal{E}^m$. It follows that $\mathcal{W}^m \setminus \mathcal{S}$ has zero Lebesgue measure. As $\mathcal{W}^m \subseteq \mathcal{V}^m$ and $\mathcal{V}^m \setminus \mathcal{V}^m$ has zero Lebesgue measure, $\mathcal{V}^m \setminus \mathcal{S}$ also has zero Lebesgue measure.\footnote{A simpler proof of the same spirit observes directly that $\mathcal{S}$ can be derived from $\mathcal{V}^m$ by excluding certain hyperplanes in $\mathcal{E}^m$. However, it is conceptually clearer to maintain the distinction between the generic conditions imposed on valuations and those imposed on valuation profiles.}

4.5.2. **Proof that in Proposition 2, 3b implies 3a**

We now establish that given the set $\mathcal{S}$ defined in (38), condition 3b of Proposition 2 implies 3a using several lemmas. Lemma 1 holds for all valuation profiles in $\mathcal{V}^m$ (not just $\mathcal{S}$), and indeed beyond $\mathcal{V}^m$ as virtually no assumption on qualitative properties of valuations is required.

**Lemma 1. Condition 3b of Proposition 2 implies (32).**

**Proof.** Assume 3b holds for consumer $i$. Choose $z \in \{0, 1\}^n$, $j, j' \in J_0(z)$ with $j \neq j'$. Then:

\[ p(z, +) + p(z + e^j, +) + p(z + e^{j'}, +) = v_i(z + e^j) - v_i(z) + v_i(z + e^{j'}) - v_i(z) = p(z, +) + p(z + e^j) + p(z + e^{j'}) - p(z). \]

**Lemma 2. For all valuation profiles in $\mathcal{S}$, whenever condition 3b of Proposition 2 holds for consumer $i$:**

\[
\begin{align*}
D_i^j(p(z, +)) &= z + V_0(z), \quad \forall z \in \{0, 1\}^n \setminus 1, \\
D_i^j(p(z, -)) &= z - V_1(z), \quad \forall z \in \{0, 1\}^n \setminus 0.
\end{align*}
\]

**Proof.** Consider valuation profile $v \in \mathcal{S}$. Assume 3b holds for consumer $i$. Strict monotonicity (see Observation 1) and the normalization $v_i(0) = 0$ (see Assumption 1) imply that for all $j$, $v_i(e^j) > 0$. That is, $i$ assigns a strictly positive value to each good. It follows that because $p(z, +)$ assigns each good $j$ with $z_j > 0$ a price of zero, at $p(z, +)$, any package $z'$ demanded by a consumer with valuation in $\mathcal{W}$ must contain $z$. Symbolically: $z' \in D_i^j(p(z, +)) \Rightarrow z \leq z'$. Choose $z' \in \{0, 1\}^n$ with $z \leq z'$ and $z' \notin z + V_0(z)$. Let $K$ be the set of goods in $z'$ but not in $z$, so that $K = \{ j : z_j < z'_j \}$. For the assumed relations of $z$ to $z'$ to hold, $K$ must contain at least two elements. So if we write $K$ as a list, $K = \{ j_1, \ldots, j_k \}$, we must have $k \geq 2$. Using a telescoping sum, distributing terms, and plugging in for prices $p(z, +)_{j_k}$ from (33), we get:

\[
v_i(z') = \sum_{h=1}^k p(z, +)_{j_h} = v_i(z) + \sum_{h=1}^k \left[ v\left( z + \sum_{\ell=1}^h e^{j_\ell} \right) - v\left( z + \sum_{\ell=1}^{h-1} e^{j_\ell} \right) \right]_{j_h} = v_i(z) + \sum_{h=1}^k \left[ v\left( z + \sum_{\ell=1}^h e^{j_\ell} \right) - v\left( z + \sum_{\ell=1}^{h-1} e^{j_\ell} \right) \right]_{j_h}.
\]
\[= v_i(z) + \sum_{h=1}^{k} \left[ \left( v_i \left( z + \sum_{\ell=1}^{h} e^{\ell t} \right) - v_i \left( z + \sum_{\ell=1}^{h-1} e^{\ell t} \right) \right) - \left( v_i(z + e^{h t}) - v_i(z) \right) \right]. \quad (43)\]

Recalling that \( k \geq 2 \), strict submodularity implies that for \( h = 2, \ldots, k \):

\[v_i \left( z + \sum_{\ell=1}^{h} e^{\ell t} \right) - v_i \left( z + \sum_{\ell=1}^{h-1} e^{\ell t} \right) < v_i(z + e^{h t}) - v_i(z). \quad (44)\]

Noting that for \( h = 1 \), the inequality in (44) is replaced by an equality, (43)–(44) imply that:

\[v_i(z) > v_i(z') - \sum_{h=1}^{k} p(z, +) e^{j h}. \quad (45)\]

Recalling that all goods in \( z \) are also in \( z' \) and that the prices on the right hand side of (45) are precisely the prices of goods in \( z' \) but not in \( z \), (45) is equivalent to: \( v_i(z'|p(z, +)) > v_i(z'|p(z, +)) \). So \( z' \notin D_i(p(z, +)) \). Eq. (33) implies that at price vector \( p(z, +) \), \( i \) is indifferent between \( z \) and \( z + e^j \) for all \( j \in J_i(z) \). Since we have established above that all other bundles are inferior to \( z \) at these prices, (41) now follows. A similar argument establishes (42). \( \square \)

**Lemma 3.** For all valuation profiles in \( S \), whenever (33) in condition 3b of Proposition 2 holds for any consumer \( i \), relating \( i \)'s marginal values to price vectors \( P_+ \) and \( P_- \), all other consumers will have single-valued demand at price vectors \( P_+ \) and \( P_- \).

**Proof.** Given strict monotonicity (see Observation 1) and (27), it can never be optimal for a consumer to purchase a package not containing all goods in \( z \) at price vector \( p(z, +) \) in \( P_+ \). Eq. (36) implies that when price vectors in \( P_+ \) satisfy (33) for agent \( i \), no agent other than \( i \) will be indifferent among any pair of consumption bundles containing all goods in \( z \) at price vector \( p(z, +) \) in \( P_+ \), and hence all other agents will have single-valued demand at such price vectors. Given (28)–(29), a similar argument extends to \( P_- \) using (37). \( \square \)

Lemmas 1–3 imply that within \( S \), condition 3b of Proposition 2 implies 3a. Notice in particular that in condition 3a, \( y(z, +) \) (resp., \( y(z, -) \)) is the sum of \( z \) and the demands of agents other than \( i \) at \( p(z, +) \) (resp., \( p(z, -) \)), where these demands are guaranteed to be single-valued by Lemma 3.

### 4.5.3. Proof that in Proposition 2, 3a implies 3b.

We establish that when \( S \) is defined by (38), 3a implies 3b. Below, only Lemma 6 requires appeal to \( S \).

**Lemma 4.** If \( z \in \{0, 1\}^n \setminus \mathbf{1} \) and demand satisfies (30) for \( z \), then there is exactly one consumer with multi-valued demand at price vector \( p(z, +) \). Similarly, if \( z \in \{0, 1\}^n \setminus \mathbf{0} \) and demand satisfies (31) for \( z \), then there is exactly one consumer with multi-valued demand at \( p(z, -) \).

**Proof.** Assume demand satisfies (30) for \( z \in \{0, 1\}^n \setminus \mathbf{1} \). Let \( y = y(z, +) \) and \( y' = y(z, +) + e^{j'} \) for some \( j' \in V_0(z) \). Then \( \sum_{j=1}^{n} y_j < \sum_{j=1}^{n} y'_j \): the total number of units summed across all goods is greater in \( y' \) than \( y \). So there exists consumer \( i \) and bundles \( x, x' \in D_i(p(z, +)) \) such
that the total number of units of all goods differs in \( x \) and \( x' \). If not, total units summed across all goods would be constant over all bundles in aggregate demand \( D(p(z, +)) \), contrary to the fact by (30), \( y, y' \in D(p(z, +)) \). As total units summed across all goods differs in \( x \) and \( x' \), assume wlog:
\[
\sum_{j=1}^{n} x_j < \sum_{j=1}^{n} x'_j.
\] (46)

We have established that there is at least one consumer – namely, \( i - \) with multi-valued demand at \( p(z, +) \). Next we argue that there is at most one consumer with multi-valued demand at \( p(z, +) \). Assume for contradiction that some other consumer \( i' \), distinct from \( i \), has multi-valued demand at \( p(z, +) \). There are two distinct bundles \( w, w' \in D^{i'}(p(z, +)) \). Assume wlog:
\[
\sum_{j=1}^{n} w_j \leq \sum_{j=1}^{n} w'_j.
\] (47)

Choose any bundle \( u \) in the aggregate demand of all agents other than \( i \) and \( i' \). That is, \( u \in \sum_{i'' \in \{i, i'\}} D^{i''}(p(z, +)) \). The above implies that the three bundles \( t, t' \) and \( t'' \) defined by: \( t = u + x + w, \ t' = u + x + w', \) and \( t'' = u + x' + w' \), are distinct, and \( t, t' \), and \( t'' \) all belong to \( D(p(z, +)) \). Conditions (46)–(47) imply: \( \sum_{j=1}^{n} t_j < \sum_{j=1}^{n} t'_j \) and \( \sum_{j=1}^{n} t_j < \sum_{j=1}^{n} t''_j \). So the two distinct bundles, \( t \) and \( t' \), both contain fewer total units of all goods than the bundle \( t'' \). But this contradicts the fact that if \( D(p(z, +)) \) has structure (30), then all bundles but one in \( D(p(z, +)) \) contain the same total number of units summed across all goods and the remaining bundle contains fewer units. As \( i \neq i' \) leads to a contradiction, it must be that \( i = i' \), so that exactly one consumer has multi-valued demand at \( p(z, +) \). A similar argument shows that for \( z \in \{0, 1\}^n \setminus 0 \), there exists exactly one consumer with multi-valued demand at \( p(z, -) \). □

**Lemma 5.** If \( z \in \{0, 1\}^n \setminus 1 \) and demand satisfies (30) for \( z \), then there exists a consumer \( i = i(z, +) \) satisfying:
\[
p(z, +)_j = v_i(z + e^j) - v_i(z), \quad \forall j \in J_0(z).
\] (48)

Similarly, if \( z \in \{0, 1\}^n \setminus 0 \) and demand satisfies (31) for \( z \), then there exists a consumer \( i' = i(z, -) \) satisfying:
\[
p(z, -)_j = v_i'(z) - v_i'(z - e^j), \quad \forall j \in J_1(z).
\] (49)

**Proof.** Lemma 4 implies that for all \( z \in \{0, 1\}^n \setminus 1 \), if demand satisfies (30) for \( z \), then there exists a consumer \( i = i(z, +) \) and bundle \( z' \) such that
\[
D^i(p(z, +)) = \{ z' \} \cup \{ z' + e^j : j \in J_0(z) \}.
\] (50)

It must be that \( z' \leq z \) because if \( 1 = z'_j > z_j = 0 \), then for some \( j \in J \) (so that \( j \in J_0(z) \)), \( z' + e^j \) belongs to \( D^i(p(z, +)) \) but not \( \{0, 1\}^n \), contradicting \( D^i(p(z, +)) \subseteq \{0, 1\}^n \), which is assumed under 1UPG (see (7) and (9)). Assume for contradiction that \( z' < z \). Then there is \( j \in J \) such that \( z_j = 1 \) and \( z'_j = 0 \). Then by (27), \( p(z, +)_j = 0 \), implying not only do we have \( z' \in D^i(p(z, +)) \), but also: \( z'_j + e^j \in D^i(p(z, +)) \) where \( j \notin J_0(z) \). This contradicts (50). So \( z' = z \). Eq. (50) implies that \( i \) is indifferent between \( z \) and \( z + e^j \) for all \( j \in J_0(z) \), which is equivalent to (48). A similar argument establishes the second statement of the lemma. □
Lemma 6. For all valuation profiles in \( S \), whenever condition 3a of Proposition 2 holds, there exists a unique consumer \( i^* \) such that for all \( z \in \{0, 1\}^n \setminus 1 \),

\[
p(z, +)_j = v_{i^*}(z + e^j) - v_{i^*}(z), \quad \forall j \in J_0(z),
\]

and for all \( z \in \{0, 1\}^n \setminus 0 \),

\[
p(z, -)_j = v_{i^*}(z) - v_{i^*}(z - e^j), \quad \forall j \in J_1(z).
\]

Equivalently, within \( S \), we can find a single (unique) consumer \( i^* \) such that:

\[
i^* = i(z, +) = i(z', -), \quad z \in \{0, 1\}^n \setminus 1, \forall z' \in \{0, 1\}^n \setminus 0.
\]

where \( i(z, +) \) and \( i(z, -) \) are defined by the statement of Lemma 5.\(^6\)

Whereas Lemma 5 found for each bundle \( z \), an agent \( i(z, +) \) satisfying (48) and also an agent \( i(z, -) \) satisfying (49), Lemma 6 reverses the order of the quantifiers \( \forall \) and \( \exists \), and finds a single agent \( i^* \), who, for all bundles \( z \) satisfies both (48) and (49).

Proof of Lemma 6. If we find \( i^* \) satisfying (51)–(52), then (39) of Proposition 3 implies that such an \( i^* \) must be unique. Lemma 5 implies that there is a consumer \( i(0, +) \) satisfying (48) for \( z = 0 \). Define \( i^* := i(0, +) \). We will argue that within \( S \), \( i^* \) satisfies (53), or equivalently, (51)–(52). We establish Lemma 6 inductively, where the induction is on the number \( \sum_{j=1}^n z_j \) of goods in package \( z \). The argument is split into a base case and an inductive step.

Base case: \( b_1 \) \( i^* \) satisfies (51) when \( \sum_{j=1}^n z_j = 0 \). \( b_2 \) \( i^* \) satisfies (52) when \( \sum_{j=1}^n z_j = 1 \).

Inductive step:

\( i_1 \) If \( i^* \) satisfies (51) for all \( z \) such that \( \sum_{j=1}^n z_j = k \), then \( i^* \) satisfies (52) for all \( z \) such that \( \sum_{j=1}^n z_j = k + 2 \).

\( i_2 \) If \( i^* \) satisfies (52) for all \( z \) such that \( \sum_{j=1}^n z_j = k + 1 \), then \( i^* \) satisfies (51) for all \( z \) such that \( \sum_{j=1}^n z_j = k \).

If we establish the base case and the inductive step, then the lemma will follow. To see this, let \( k = \sum_{j=1}^n z_j \) be the number of goods in package \( z \). Then, having established the base case, we have established that \( i^* \) satisfies (51) when \( k = 0 \) and \( i^* \) satisfies (52) when \( k = 1 \). Then applying \( i_1 \), it follows that \( i^* \) satisfies (52) when \( k = 2 \), and applying \( i_2 \), it follows that \( i^* \) satisfies (51) for \( k = 1 \), applying \( i_1 \), \( i^* \) satisfies (52) when \( k = 3 \), applying \( i_2 \), \( i^* \) satisfies (51) when \( k = 2 \), and so on, until we establish that \( i^* \) satisfies (52) when \( k = n \) and \( i^* \) satisfies (51) when \( k = n - 1 \), and we are done.

For the base case, \( b_1 \) is an immediate consequence of the fact that we have set \( i^* = i(0, +) \) and Lemma 5. Next, consider \( z \) for which \( \sum_{j=1}^n z_j = 1 \). Then \( z = e^j \) for some \( j \in J \), and we have:

\[
p(e^j, -)_j = p(0, +)_j = v_{i^*}(0 + e^j) - v_{i^*}(0) = v_{i^*}(e^j) - v_{i^*}(e^j - e^j), \quad \text{where the first equality follows from (29), and the second from } b_1, \text{ which was established above. This establishes } b_2.
\]

We now perform the inductive step. First, we establish \( i_1 \). Suppose \( i^* \) satisfies (51) for all \( z \in \{0, 1\}^n \) with \( \sum_{j=1}^n z_j = k \), and consider \( z' \in \{0, 1\}^n \) with \( \sum_{j=1}^n z'_j = k + 2 \). Then there exists

\(^6\) Note that Lemma 4 implies that the consumers \( i(z, +) \) and \( i(z, -) \) from Lemma 5 are unique, and so cannot be chosen in multiple ways.

\(^7\) Note that (52) is not required when \( \sum_{j=1}^n z_j = 0 \), or equivalently, when \( z = 0 \).
\[ z \in \{0, 1\}^n \text{ with } \sum_{j=1}^nz_j = k \] and \( j', j'' \in J_0(z) \) with \( j' \neq j'' \), and such that \( z' = z + e_j + e_j'' \). Since (51) holds for \( z \) with \( \sum_{j=1}^nz_j = k \),
\[ p(z, +)j' = v_{i^*}(z + e_j) - v_{i^*}(z) \quad \text{and} \quad p(z, +)j'' = v_{i^*}(z + e_j'') - v_{i^*}(z). \]
Subtracting the first equation from the second:
\[ p(z, +)j'' - p(z, +)j' = v_{i^*}(z + e_j'') - v_{i^*}(z + e_j) - v_{i^*}(z + e_j) + v_{i^*}(z + e_j''). \]
Letting, \( i = i(z, -) \), (49) implies that:
\[ p(z', -)j = v_{i^*}(z') - v_{i^*}(z' - e_j), \quad \forall j \in J_1(z'). \]
Because \( j', j'' \in J_1(z') \), and also because \( z' - e_j' = z + e_j'' \) and \( z' - e_j'' = z + e_j' \), it follows that
\[ p(z', -)j = v_{i^*}(z') - v_{i^*}(z + e_j') \quad \text{and} \quad p(z', -)j'' = v_{i^*}(z') - v_{i^*}(z + e_j''). \]
Subtracting the first equation from the second:
\[ p(z', -)j'' - p(z', -)j' = v_{i^*}(z + e_j') - v_{i^*}(z + e_j''). \]
By (32):
\[ p(z, +)j'' - p(z, +)j' = p(z + e_j'') + p(z + e_j') - p(z + e_j) - p(z + e_j''). \]
Applying (29) to the terms on the right hand side of (57) in light of the fact that \( z'(z + e_j') + e_j'' = (z + e_j) + e_j'' \) yields:
\[ p(z, +)j'' - p(z, +)j' = p(z', -)j'' - p(z', -)j'. \]
It now follows from (54), (56), and (40) that \( i = i^* \). Appealing to (55), this establishes i1.
To establish i2, suppose (52) holds for all \( z \in \{0, 1\}^n \) with \( \sum_{j=1}^nz_j = k + 1 \). Consider \( z' \in \{0, 1\}^n \) with \( \sum_{j=1}^nz''_j = k \). We may assume \( k < n \) because (51) is only required for \( k = 0, 1, \ldots, n - 1 \) (or equivalently, for packages in \( \{0, 1\}^n \setminus 1 \)). So \( J_0(z') \) is nonempty. So for some \( j' \in J_0(z') \) and \( z'' \in \{0, 1\}^n \) with \( \sum_{j=1}^nz''_j = k + 1 \), \( z' = z'' - e_j' \). By (48), for \( i = i(z', +) \):
\[ p(z', +)j = v_{i^*}(z' + e_j) - v_{i^*}(z'), \quad \forall j \in J_0(z'). \]
So in particular,
\[ v_{i^*}(z' + e_j') - v_{i^*}(z') = p(z', +)j' = p(z', -)j = v_{i^*}(z'') - v_{i^*}(z'' - e_j') = v_{i^*}(z' + e_j') - v_{i^*}(z'), \]
where the second equality follows from (29) and the third equality follows from the inductive hypothesis that \( i^* \) satisfies (52) for packages \( z \) with \( \sum_{j=1}^nz_j = k + 1 \). It follows from (39) of Proposition 3 that \( i = i^* \), which together with (58) establishes i2, completing the proof. \( \square \)

Lemma 6 directly implies that within \( S \), condition 3a of Proposition 2 implies 3b.
5. Identification with multi-unit demand

The previous section established identification under the assumption that each consumer desires at most one unit of each good (1UPG). This section relaxes that assumption, allowing for multi-unit demand of each good. There is a trade-off however. To derive an identification result analogous to Theorem 1, we must employ a stronger substitutes assumption than submodularity. With this in mind, in the analysis that follows, we maintain Assumption 1, so that we continue to assume normalization and monotonicity, but we replace Assumption 2 with:

**Assumption 3** ($M^2$-concavity). For all $z, z' \in \mathbb{B}$ and $j \in J$ with $z_j > z'_j$:

$$v_i(z) + v_i(z') \leq \max \left[ v_i(z - e^j) + v_i(z' + e^j), \max_{j' : z_j > z_{j'}} \{ v_i(z - e^j + e^{j'}) + v_i(z' + e^j - e^{j'}) \} \right].$$

Unlike Assumption 2, Assumption 3 omits 1UPG; we now allow satiation points $N_j$ to be arbitrary positive integers (see Section 2.1.1). Satiation points $N_j$ are assumed common across consumers and known to the econometrician. The Online Appendix explains how we can relax this assumption (and the case $N_j = \infty$). So assume a fixed profile ($N_j : j \in J$) of satiation points. $\mathcal{V}$ denotes the set of valuations satisfying Assumptions 1 and 3 for this profile.

$M^2$-concavity (read “$M$-natural concavity”) was introduced by Murota and Shioura [26] and extensively studied by Murota [25]. $M^2$-concavity says that if bundle $z$ contains more units of good $j$ than does $z'$, the sum of values of the bundles increases either when (i) a unit of $j$ is transferred from bundle $z$ to $z'$, or (ii) a unit of good $j$ is transferred from $z$ to bundle $z'$ and, in exchange, a unit of some good $j'$ – of which $z'$ contains more units than $z$ – is transferred from $z'$ to $z$. For this reason, $M^2$-concavity is often referred to as an exchange axiom.

Murota and Tamura [28] showed that among functions on the discrete grid $\mathbb{Z}_+^n$ (or appropriate subsets thereof) which are extendible to concave functions on $\mathbb{R}_+^n$, $M^2$-concavity is equivalent to the well known gross substitutes property (see Section 3.2).\(^8\) The $M^2$-concave valuations are equivalent to the strong substitutes valuations of Milgrom and Strulovici [23],\(^9\) where the latter are defined to be valuations satisfying the gross substitutes property under non-linear prices, when each unit of each good is priced separately. That $M^2$-concave valuations are so closely allied to the well known property of gross substitutes motivates their study. However, in contrast to the gross substitutes condition, which defines such valuations indirectly in terms of comparative statics of demand, the $M^2$-concavity formulation defines these valuations directly in terms of inequalities they must satisfy. This explicit characterization will be an advantage, especially when we deal with the issue of the genericity of identification.

One class of $M^2$-concave valuations (also satisfying Assumption 1) are those of the form:

$$v_i(z) := u_0 \left( \sum_{j=1}^{n} z_j \right) + \sum_{j=1}^{n} u_j(z_j), \quad \forall z \in \mathbb{B}, \quad (59)$$


\(^9\) It follows from Theorems 6.42 and 11.4 of Murota [25] and Theorem 13 of Milgrom and Strulovici [23].
where for \( j = 0, \ldots, n, u_j: \mathbb{R} \to \mathbb{R} \) is increasing and concave with \( u_j(0) = 0 \). Valuations of the form \((59)\) generalize the concave transformations of additive valuations discussed in Section 3.2. For further examples of \( M^2 \)-concave valuations, see Section 6.3 of Murota [25].

Proposition 4 presents properties of \( M^2 \)-concave functions that are useful for Theorem 2.

**Proposition 4.** Let \( v_j: \mathbb{B} \to \mathbb{R} \) be an \( M^2 \)-concave valuation. Then,

1. \( v_i \) is submodular: \( v_i(z + e^j + e^\ell) - v_i(z + e^\ell) \leq v_i(z + e^j) - v_i(z), \forall z \in \mathbb{B}, \forall j, \ell \in J \setminus J_s(z) \) with \( j \neq \ell \).
2. \( v_i \) is concave-extendible, meaning that \( v_i \) can be extended to a concave function on \( \mathbb{R}^n \). So \( v_i \) is component-wise concave: \( v_i(z + 2e^j) - v_i(z + e^j) \leq v_i(z + e^j) - v_i(z), \forall z \in \mathbb{B}, \forall j \in J_{s-2}(z) \).
3. \( v_i \) satisfies strong self-substitutability: \( v_i(z + 2e^j) - v_i(z + e^j) \leq v_i(z + e^j + e^\ell) - v_i(z + e^\ell), \forall z \in \mathbb{B}, \forall j \in J_{s-2}(z), \forall \ell \in J \setminus J_s(z) \) with \( j \neq \ell \).
4. \( v_i \) satisfies the gross substitutes property, meaning that when demand \( D^i \) is generated by valuation \( v_i: \forall p, p' \in \mathbb{R}^n_+, p \leq p' \Rightarrow \forall z \in D^i(p), \exists z^i j \in D^i(p') \forall j \in J, p_j = p'_j \Rightarrow z_j \leq z^i_j \).

Part 3 is proven in the Online Appendix.\(^{10}\) All other parts of Proposition 4 can be found in Murota [25].\(^{11}\) \( M^2 \)-concavity not only implies that different goods are substitutes for one another, as illustrated by submodularity and gross substitutes, but also that different units of the same good are substitutes for one another, as illustrated by component-wise concavity. Whereas submodularity says that the incremental value of good \( j \) diminishes when another good \( \ell \) is added to the bundle, component-wise concavity says that the incremental value of good \( j \) diminishes when another unit of good \( j \) is added to the bundle. Strong self-substitutability compares these two effects: It says that adding another unit of good \( j \) decreases the incremental value of good \( j \) by more than adding a unit of another unit of good \( \ell \).

Under 1UPG, Lehmann, Lehmann, and Nisan [18] showed that unlike the submodular valuations, the gross substitutes valuations have Lebesgue measure zero in \( E \) (see Section 3.3). A similar result applies when we allow multi-unit demand. As \( M^2 \)-concavity implies gross substitutes, this establishes a sense in which our identification result with multi-unit demand requires a stronger substitutes assumption than was required for identification under 1UPG.

The proof of Theorem 1 depended on the fact that set of valuations satisfying Assumptions 1 and 2 had positive Lebesgue measure in the ambient Euclidean space. The fact that the set of \( M^2 \)-concave valuations, and hence the set \( \tilde{V} \) of valuations satisfying Assumption 1 and 3, has zero Lebesgue measure in \( E \) complicates the formulation and proof of *genericity* of identification. A natural response would be to treat the ambient Euclidean space as the affine hull of \( \tilde{V} \) within \( E \), denoted aff(\( \tilde{V} \)), rather than as \( E \) itself. However, as is shown in the Online Appendix, the set of \( M^2 \)-concave valuations – and also the set \( \tilde{V} \) – is not a convex set. So it is possible that the dimension of aff(\( \tilde{V} \)) is greater than the dimension of \( \tilde{V} \), which would imply that \( \tilde{V} \) has zero Lebesgue measure in aff(\( \tilde{V} \)). Proposition 5 helps break the impasse.

\(^{10}\) Lemma 1 of Milgrom and Strulovic [23] establishes the property we call “strong self-substitutability” for the strong substitutes valuations. Because, as explained below, the strong substitutes valuations can be shown to be the same as the \( M^2 \)-concave valuations, this provides an alternative proof of part 3 of Proposition 4.

\(^{11}\) See Theorems 6.19, 6.42 and 11.5 of Murota [25], originally proven by Murota and Shioura [27], Murota [24], Danilov, Koshevoy, and Lang [7], Fujishige and Yang [10], and Murota and Tamura [28].
Proposition 5. There exists a finite collection \( \{ V_f : f \in F \} \) of sets (where \( F \) is an index set), such that each set \( V_f \) is a convex subset of \( E \), and \( \tilde{V} = \bigcup_{f \in F} V_f \). Consequently, letting \( G \) be the set of functions from the set of agents \( I \) to the index set \( F \), the set of valuation profiles can also be written as a finite union of convex sets:

\[
\tilde{V}^m = \bigcup_{g \in G} [\times_{i \in I} V^*_g(i)] =: \bigcup_{g \in G} V^*_g. \tag{60}
\]

Section VI.1 of the Online Appendix defines the sets \( V_f \) and proves Proposition 5. A definition of the sets \( V_f \) also provides a definition of the sets \( V^*_g \) via Eq. (60): that is, \( V^*_g := \times_{i \in I} V^*_g(i) \) for all \( g \in G \).

\( \tilde{V} \) is not convex because the \( M^2 \)-convexity inequalities contain a maximization operator. Each such inequality says that one of several linear inequalities\(^\text{12} \) holds; for each \( z, z' \in \mathbb{B} \), and \( j \in J \) with \( z_j > z'_j \), the inequality required by \( M^2 \)-concavity can be rewritten as a disjunction:

\[
\begin{align*}
\text{either} & \quad v_i(z) + v_i(z') \leq v_i(z - e^j) + v_i(z' + e^j), \\
\text{or} & \quad v_i(z) + v_i(z') \leq v_i(z - e^j + e^j_i) + v_i(z' + e^j_j - e^j_i), \\
\text{or} & \quad v_i(z) + v_i(z') \leq v_i(z - e^j + e^j_i) + v_i(z' + e^j_j - e^j_i) + v_i(z' + e^j_j - e^j_i), \\
\text{or} & \quad \vdots \\
\text{or} & \quad v_i(z) + v_i(z') \leq v_i(z - e^j + e^j_i) + v_i(z' + e^j_j - e^j_i),
\end{align*}
\tag{61}
\]

where \( \{ j_1, \ldots, j_k \} \) is the set of good indices \( j_h \) such that \( z'_{j_h} > z_j \). If for each pair \( z, z' \) and each \( j \) with \( z_j > z'_j \), we select one inequality in (61), the combination of selected inequalities defines a polyhedron in \( E \). Each polyhedron is convex; the union of these polyhedra equals the set of \( M^2 \)-concave valuations. (Section VI.1 of the Online Appendix defines the sets \( V_f \) similarly but slightly differently.)

We now present a definition that is analogous to Definition 1 in Section 2.2.2, but accounts for the fact that the set of valuations \( \tilde{V} \) is neither full dimensional in \( E \) nor convex.

Definition 2. Call an element \( V^*_g \) of \( \{ V^*_g : g \in G \} \) a component of the set \( \tilde{V}^m \) (see Proposition 5). Let \( \mathcal{X} \) be the set of all valuation profiles \( v \in \tilde{V}^m \) such that valuations are not identified from demand. Then valuations are component-wise generically identified from demand if for each \( g \in G \), \( V^*_g \cap \mathcal{X} \) has zero Lebesgue measure in the affine hull of \( V^*_g \) (denoted \( \text{aff}(V^*_g) \)).

To grasp the significance of this definition, note that whereas each set \( V^*_g \) has zero Lebesgue measure in \( E^m \), \( V^*_g \) has infinite Lebesgue measure in \( \text{aff}(V^*_g) \). To be precise, \( \text{aff}(V^*_g) \) is an affine space, and as it contains the origin, it is in fact a subspace of \( E^m \), and so \( \text{aff}(V^*_g) \) is isomorphic to a Euclidean space. So we can define Lebesgue measure on \( \text{aff}(V^*_g) \). Because \( V^*_g \) is convex, and hence has nonempty relative interior (i.e., nonempty interior relative to its affine hull \( \text{aff}(V^*_g) \)), and any measurable set with nonempty interior in a Euclidean space has nonzero Lebesgue measure in that space, \( V^*_g \) has nonzero Lebesgue measure in \( \text{aff}(V^*_g) \). Moreover the fact that \( V^*_g \) is a cone can be used to show that not only does \( V^*_g \) have positive Lebesgue measure in \( \text{aff}(V^*_g) \), but \( V^*_g \) has infinite Lebesgue measure in \( \text{aff}(V^*_g) \).

\(^{12}\) To be precise, the inequalities are linear in valuations conceived of as points in \( E \).
The genericity notion for single-unit demand implied identification with probability 1 if valuation profiles are selected from a probability measure absolutely continuous with respect to Lebesgue measure (see Section 2.2.2). Component-wise generic identification warrants a similar but more involved statement. Let \( \mu \) be any probability measure over valuation profiles in \( \mathcal{V}^m \). For any \( X \subseteq \mathcal{V}^m \) with \( \mu(X) > 0 \), let \( \mu^X \) be the conditional probability measure resulting from conditioning \( \mu \) on \( X \). Assume that \( \mu \) is such that for all \( g \in G \), if \( \mu(\mathcal{V}^*_g) > 0 \), then \( \mu^g \) is absolutely continuous with respect to Lebesgue measure on \( \text{aff}(\mathcal{V}^*_g) \). Then if valuation profiles are selected according to \( \mu \), with probability 1, valuations will be identified from demand.

We now present our identification result for multi-unit demand.

**Theorem 2.** Suppose there are at least two goods \((n \geq 2)\). Then, under Assumptions 1 and 3, valuations are component-wise generically identified from demand.

For the proof, see Appendix A below as well as the Online Appendix.

6. Discussion

6.1. Relation to McManus [22]

McManus [22] studies a generic identification property in simultaneous equations of parametric models, providing conditions on the relations between the dimensions of the endogenous variables and parameters for what he referred to as generic identification. McManus studies models of the form \( Y_t = f(X_t, U_t, \theta) \), where \( Y_t \) is the set of endogenous variables, \( f(\cdot) \) is the system of equations, \( X_t \) is the set of exogenous variables, \( \theta \) is the parameter of interest, and \( U_t \) contains unobserved disturbance terms. A fundamental difference between our paper and McManus is that by generic identification, McManus means that for almost all models \( f \), identification holds everywhere under \( f \), that is, for all parameter values. In contrast, we study a specific model and show that in our fixed model, we have identification for almost all parameter values, where for us the parameter values are the valuation profiles assigning a value to each bundle for each consumer. Therefore the arguments establishing genericity of identification are very different. In McManus’s problem, the genericity simply follows from the Whitney embedding theorem but in our problem we have to develop our own results. There are various other differences such as the fact that we deal with a discrete setting in which demand changes discontinuously when some consumer adds or removes bundles from her demand in response to a price change, and that we study a measure-theoretic notion of genericity as opposed to McManus’s topological notion. On the most fundamental level, our identification results rely on qualitative assumptions on valuations relating to the substitutes properties of submodularity and \( M^2 \)-concavity. One can show, for example, that in the case of multi-unit demand, if valuations were strictly component-wise convex – which is inconsistent with \( M^2 \)-concavity – individual demand for a single consumer would fail to be identified from prices, and so a fortiori, aggregate demand would not be identified. It is difficult to see how arguments similar of the McManus’s general topological arguments could deliver this difference between the case where different units of the same good are complements just described and where generic identification fails and the case of substitutes we study where it succeeds. Finally our proofs, unlike those of McManus, are constructive.

---

13 The Online Appendix shows that with only one good \((n = 1)\), when consumers may demand many units, valuations are not generically identified from demand.
6.2. Relation to the econometric literature on discrete choice

Our framework allows us to attain much stronger identification results than have been previously found in the literature of discrete choice. However, our current model is less adapted to empirical applications. Our model differs from those employed in the empirical literature on discrete choice in several ways. First, we assume a finite population of consumers rather than a continuum. Second, our model does not have a stochastic component in the utility such as logit error. A more superficial difference is that we do not employ a characteristics-based approach. The advantage of our model is that we employ a rich model of consumer preferences taken from literatures on Walrasian equilibrium with discrete goods and combinatorial auctions. In view of these differences, it is natural to ask about the relevance of our results for econometrics.

It is a challenging but important task to integrate models of rich preferences, such as those that we study into discrete choice and econometrics. We do not expect our results to translate directly once such an approach has been developed, but we do expect that our results will provide intuitions about the sort of analogous results that one might expect to obtain in a stochastic setting. In particular, we believe that the closest analogs of our identification results that one can expect to obtain in such settings will be partial identification results. However there will be several similarities: the same qualitative assumptions – such as submodularity for single unit demand and $M^2$-concavity for multi-unit demand – that we found useful for our identification results will be relevant for the partial identification results that one is likely to be able to obtain in the stochastic case and this will be so for similar reasons and using related arguments. As in our model one of the main challenges facing the econometrician would be that the econometrician only observes marginal shares (i.e., the demand for each good demanded individually) but might like to make inferences about the distribution of shares of packages. We believe that arguments related to the ones that we employ can be used to bound the shares of packages using the observed shares of individual goods, and to bound changes in shares of bundles with changes in price. In an unpublished note (Kim and Sher [17]), which is available from the authors upon request, we make some tentative suggestions along these lines. How far one could go toward point identification in such an approach is an open question that should be addressed with further research. Of course, many other issues would arise in implementing a model related to ours in an econometric discrete choice setting, such as the price endogeneity problem, but one could try to address these along standard lines.

6.3. Extensions

Here we briefly mention some extensions of our results. These extensions are discussed in greater detail in the Online Appendix. Our multi-unit demand identification result incorporates a bound on the number of units of each good demanded by each consumer. With unbounded demand, we can establish identification under conditions analogous to the case of bounded multi-unit demand, but have not yet established that these analogous conditions are generic relative to our qualitative assumptions on valuations. Restricting attention to bounded multi-unit demand, we assumed above that the satiation points $N_j$ for each good are known and common across consumers. In the Online Appendix, we explain how, even if these satiation points were unknown to the econometrician and heterogenous across consumers, they could be identified (up to permutations of the consumer indices). Similarly, the econometrician can identify the number of consumers if this is not initially known. Finally, our identification results would continue to hold if instead of a finite number of consumers, there were a continuum of consumers partitioned into
a finite number of types, possibly differing in their mass, such that all consumers of a given type share the same valuation.

Appendix A. A result that implies Theorem 2

The proof of Theorem 2 parallels that of Theorem 1 but is more involved. Analogously to the proof of Theorem 1, we proceed by presenting a result that implies Theorem 2. Here we present this result. This result (Proposition 6) is proven in the Online Appendix. The argument that to prove Theorem 2, it is in fact sufficient to prove Proposition 6 is similar to the argument of Section 4.5 showing that Proposition 2 is sufficient for Theorem 1, and is omitted.

Proposition 6. Suppose there are at least two goods \((n \geq 2)\). Suppose, moreover, that Assumptions 1 and 3 hold. There exists a set of valuation profiles \(\hat{S} \subseteq \hat{V}^m\) with the following properties:

1. If \(v = (v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) \in \hat{S}\), where \(i, j' \in I\) are such that \(i \neq i'\), then \(v_i = v_{j'}\).
2. \(V_g^* \setminus \hat{S}\) has Lebesgue measure zero in \(\text{aff}(V_g^*)\) for all components \(V_g^*\) of \(\hat{V}_g^*\) (Section VI.1 of the Online Appendix provides a formal definition of the components \(V_g^*\) of \(\hat{V}\); see Section 5 for a discussion.)
3. Choose any valuation profile \(v = (v_1, \ldots, v_i, \ldots, v_n) \in \hat{S}\). Let \(P_+ = \{p(z, +) : z \in \mathbb{B}\}\) and \(P_- = \{p(z, -) : z \in \mathbb{B}\}\) be two collections of price vectors satisfying:

\[
p(z, +)_j = 0, \quad \forall j \in J_a(z); \quad p(z, -)_j = M, \quad \forall j \in J_0(z); \quad \text{and} \quad p(z + e^j, -)_j = p(z, +)_j, \quad \forall j \in J \setminus J_a(z),
\]

where \(M\) is a number sufficiently large that under \(v\), for every good \(j\), aggregate demand for \(j\) is always zero when the price of \(j\) is \(M\). Then the following conditions are equivalent:

(a) For all \(z \in \mathbb{B}\),

\[
\text{there exist } y(z, +) \in \mathbb{Z}^n_+ \text{ such that } D_v(p(z, +)) = y(z, +) + V_0(z), \quad \text{and (63)}
\]

\[
\text{there exist } y(z, -) \in \mathbb{Z}^n_+ \text{ such that } D_v(p(z, -)) = y(z, -) - V_1(z). \quad \text{(64)}
\]

Moreover, \(\forall z \in \mathbb{B}, \forall j, j' \in J \setminus J_a(z)\) with \(j \neq j'\):

\[
p(z, +)_j + p(z + e^j, +)_{j'} = p(z, +)_{j'} + p(z + e^{j'}, +)_{j},
\]

(b) There exists a unique consumer \(i \in I\) such that:

\[
p(z, +)_j = v_i(z + e^j) - v_i(z), \quad \forall z \in \mathbb{B}, \forall j \in J \setminus J_a(z).
\]

Appendix B. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jet.2014.07.009.

References