TESTING FOR SEPARABILITY IS HARD

FEDERICO ECHENIQUE

Abstract. This paper shows that it is computationally hard to decide (or test) whether a consumption data set is consistent with separable preferences.

1. Introduction

The assumption of separable preferences is ubiquitous in economics. Economists assume separability of preferences, virtually without ever testing this assumption empirically. Here I argue that there is a reason for such lack of empirical scrutiny: The problem of deciding if a data set is consistent with separable preferences is computationally hard. There cannot exist a test that is practical on large data sets, and that serves to test for separability.

Every empirical study on consumption assumes, explicitly or implicitly, that preferences are separable. For example, data on supermarket purchases are used in isolation from other consumption decisions. Data on consumption in one year is used without regard for intertemporal consumption decisions. The choice among different goods is analyzed while ignoring any consumption/leisure tradeoffs, and independently of the allocation of financial assets. All such analyses, which depend on certain compartmentalizations in the economy, rely on the assumption of separability. It is difficult to imagine a paper in applied economics that does not make some use of separability.

A few authors have proposed tests for separability. Varian (1983) has a test that involves solving a system of polynomial inequalities. Varian's is a test for the joint hypothesis of separability and concavity of utility. Cherchye, Demuynck, and De Rock (2011) provide a computational approach to handling Varian's system of inequalities. Quah (2012) has a test for separability alone, without the assumption of concavity of utility. Quah's test is finite, meaning that one would need to check if the data fit a finite number of configurations. These tests are

---

Division of the Humanities and Social Sciences, Caltech. I am grateful to Thomas Demuynck, Bram De Rock, and John Quah for comments on the first draft of this note.

all difficult to take to data because they are computationally inefficient, and may be infeasible in large datasets. **?** test (and reject) an important implication of separability using parametric econometric methods; they do not develop a non-parametric revealed preference test.

My contribution here is to show that separability is inherently hard. Specifically, that is is NP complete. From a technical viewpoint, the contribution is to carry out a *reduction to a revealed preference graph with a fixed number of goods.* One can construct any graph as a revealed preference graph.[1] But very little is known about which graphs can be constructed with a fixed number of goods. We know, for example, that when the number of goods is two, all acyclic revealed preference graphs must be 2-acyclic; and that the complete graph is a revealed preference graph with two goods.[2] We know little else. So it is important to carry out a construction as I do here with a fixed number of goods.

A fixed number of goods is also important for the economic meaning of the result. A hardness result for testing separability appears in Cherchye, Demuynck, and De Rock (2011), but their result is asymptotic in the number of goods as well as the number of observations. Cherchye et. al. use the reduction to a revealed preference graph with as many goods as observations. For economists, this means that separability is hard if one has a large data set *and many goods.* In fact, most studies in economics use only a handful of goods. The construction used in the proof of the theorem below uses 9 goods, the same number as in the classical study on consumption by Deaton (1974). The number 9 is not a limitation of old data sets and classical studies; recent studies on consumption also use a small number of goods (for example, in their empirical analysis, Cherchye, Demuynck, and De Rock (2011) themselves use 15 goods).

Finally, I should emphasize that my result is on separability alone, with no other joint assumptions. The main motivation for using separability is two-stage budgeting. It is easy to see that two-stage budgeting follows from separability alone, and does not need convex preferences or other auxiliary assumptions. The exercise in this note is also in line with Quah's investigation of the pure content of separability. (Cherchye, Demuynck, and De Rock (2011) focus on Varian's conditions for concave separability, and it does not follow from their result that separability alone is hard to test for.)

[1]Given a directed graph $(\{1,\ldots,K\},E)$, there is a consumption dataset with K goods $\{(x^k,p^k):k=1,\ldots,K\}$ with $(k,l)\in E$ iff $x^k\neq x^l$ and $p^k\cdot x^k\geq p^k\cdot x^l$ (see **?**).

[2]Recent progress for the cases of two and three goods are in **?**

2. Testing separability

We take consumption space to be $\mathbf{R}^{n+m}_+$. The set of available goods is partitioned in two, and we write a consumption bundle as $x = (z, o) \in \mathbf{R}^{n+m}_+$. There are n goods of "type z" and m of "type o."

A *data set* is a collection (x_k, p_k), $k = 1, \ldots, K$ in which for every k $x_k \in \mathbf{R}^{n+m}_+$ is a consumption bundle purchased at prices $p_k \in \mathbf{R}^{n+m}_{++}$. Let $x_k = (z_k, o_k) \in \mathbf{R}^n_+ \times \mathbf{R}^m_+$.

A data set (x_k, p_k), $k = 1, \ldots, K$ is *rationalizable by separable preferences* if there are monotone increasing functions $u : \mathbf{R}^n_+ \to \mathbf{R}$ and $v : \mathbf{R}^{1+m} \to \mathbf{R}$ such that

$$v(u(z), o) < v(u(z_k), o_k)$$

for all $(z, o) \in \mathbf{R}^{n+m}_+$ with $p^z_k \cdot z + p^o_k \cdot o \leq p^z_k \cdot z_k + p^o_k \cdot o_k$ and $(z, o) \neq (z_k, o_k)$.

Theorem. *The problem of deciding if a dataset has a separable rationalization is NP-complete.*

3. Proof

I shall not prove that the problem is in NP. This follows, for example, from the result of Quah (2012).

3.1. Notation and definitions.

By e_i we denote the ith unit vector in $\mathbf{R}^n$, that is the vector that has zero in every entry except the ith, in which it has a one. Write e_{12} for $e_1 + e_2$. The *embedding* of $\mathbf{R}^n$ into $\mathbf{R}^m$, with $n < m$ is the function that maps a vector $x \in \mathbf{R}^n$ into the vector $(x, 0, \ldots, 0)$ in $\mathbf{R}^m$, which coincides with x_i in the first n entries and then has a zero in the remaining entries.

A *graph* is a set X together with a binary relation $R \subseteq X \times X$. We write $x \, R \, y$ for $(x, y) \in R$. A sequence $x_1, \ldots x_K$ in X is a *path* from x_1 to x_K if

$$x_K \, R \, x_{K-1} \, R \, x_{K-2} \cdots x_2 \, R \, x_1.$$

A graph, or the binary relation R, is *acyclic* if for every pair x and x', with $x \neq x'$, if there is a path from x to x' then there is no path from x' to x.

3.2. Construction.

We shall reduce from three-satisfiability. Consider a formula with L clauses, $C_1, \ldots, C_L$, involving the variables $x_1, \ldots, x_I$.

The strategy for the reduction is as follows. We introduce a pair of bundles z^1_i and z^2_i for each variable i. These bundles are not comparable by revealed preference, but they are embedded into a configuration of prices and bundles such that any rationalizing preference must reflect an assignment of truth/falsehood to each variable that makes all the clauses C_l true. This is accomplished using separability: in fact separability is crucial to make the construction work with a fixed number of goods. The bundles z_i^1 and z_i^2 live in $\mathbf{R}^2$ and they are shifted by adding different amounts of the other goods so that they can play the same role in different clauses. By separability, the comparison between z_i^1 and z_i^2 must be the same in all shifted instances.

We first (Step 1) develop the construction for a single clause. Then (Step 2) we tie the different clauses together. The formal construction follows.

For each variable x_i, define the following vectors in $\mathbf{R}^2_+$: $z_i^1 = (i, 1/i)$ and $z_i^2 = (i + 1/2, 1/(i + 1/2))$. Let $\tau : \{x_1, \ldots, x_I\} \to \{0, 1\}$ be a truthtable for the variables $x_1, \ldots, x_I$. Define a binary relation B on $\{z_i^q : q = 1, 2; i = 1, \ldots, I\}$ from τ by $z_i^1 \; B \; z_i^2$ if $\tau(x_i) = 1$ and $z_i^2 \; B \; z_i^1$ if $\tau(x_i) = 0$.

Note that each of the vectors in $\{z_i^q : q = 1, 2; i = 1, \ldots, K\}$ lie on the boundary of the convex set $A = \{(\theta_1, \theta_2) : \theta_2 \geq 1/\theta_1 \text{ and } \theta_1 > 0\}$. Define p_i^q to be such that the hyperplane

$$\{(\theta_1, \theta_2) : p_i^q \cdot (\theta_1, \theta_2) = 1\}$$

supports A at z_i^q. (A simple calculation reveals that $p_i^1 = (1/(2i), i/2)$ and $p_i^2 = (1/(2i + 1)), i/2 + 1/4)$, but this does not play a role in the sequel.)

As a consequence of these definitions, we obtain the following

Lemma 1. *For all $i, i' = 1, \ldots, I$, and all $q, q' = 1, 2$, we have*

$$p_i^q \cdot z_i^q = 1 < p_i^q \cdot z_{i'}^{q'}$$

when $i' \neq i$ and/or $q' \neq q$.

To make the sequel easier to follow, we write $\rho(z_i^q)$ for p_i^q. Define two positive numbers, ε and M as follows. Let ε be such that

$$1 < \rho(z_i^q) \cdot z_i^q + \varepsilon \rho(z_i^q) \cdot (1, 1) < \rho(z_i^q) \cdot z_{i'}^{q'} \tag{1}$$

when $i' \neq i$ and/or $q' \neq q$. In second place, let M be such that

$$\rho(z_i^q) \cdot z_{i'}^{q'} + \varepsilon \rho(z_i^q) \cdot (1, 1) < M \tag{2}$$

when $i' \neq i$ and/or $q' \neq q$.

3.2.1. *Step 1: The construction for a single clause.* Consider a single clause C. Say that $C = y_i \vee y_j \vee y_h$, with $y \in \{x, \bar{x}\}$. Let $\hat{z}_i^q = z_i^q$ and $\hat{p}_i^q = p_i^q$ if $y_i = x_i$, and $\hat{z}_i^q = z_i^{3-q}$ and $\hat{p}_i^q = p_i^{3-q}$ if $y_i = \bar{x}_i$. Define $\hat{z}_j^q$, $\hat{z}_h^q$, $\hat{p}_j^q$, and $\hat{p}_h^q$ analogously.

We are going to map the clause C into a dataset in $\mathbf{R}^8$.

Consider the observations $\{(w_i, r_i) : i = 1, 2, 4, 5, 7, 8\}$ and the set of bundles $\{w_i : i = 1, \ldots, 9\} \subset \mathbf{R}^8_+$ defined as follows:

$$
\begin{array}{ll}
w_1 = \hat{z}_i^2 + e_3 & r_1 = \rho(\hat{z}_i^2) + Me_3 + 2M(e_5 + e_6) \\
w_2 = \hat{z}_j^1 + \varepsilon(e_1 + e_2) + e_4 & r_2 = \rho(\hat{z}_j^1) \\
w_3 = \hat{z}_j^1 + e_3 + e_7 & \\
w_4 = \hat{z}_j^2 + e_3 + e_7 & r_4 = \rho(z_j^2) + Me_3 + 2M(e_4 + e_6) \\
w_5 = \hat{z}_h^1 + \varepsilon(e_1 + e_2) + e_5 & r_5 = \rho(\hat{z}_h^1) \\
w_6 = \hat{z}_h^1 + e_3 + e_8 & \\
w_7 = \hat{z}_h^2 + e_3 + e_8 & r_7 = \rho(\hat{z}_h^2) + Me_3 + 2M(e_4 + e_5) \\
w_8 = \hat{z}_i^1 + \varepsilon(e_1 + e_2) + e_6 & r_8 = \rho(\hat{z}_i^1) \\
w_9 = \hat{z}_i^1 + e_3 &
\end{array}
$$

This means that bundles w_i are purchased at prices r_i, for $i = 1, 2, 4, 5, 7, 8$. The bundles w_3, w_6, w_9 are added for convenience.

Let X be the set of products taken from the first two, and the last six, entries of the vectors w_k. That is, X is the set of pairs (z, o), with $z \in \mathbf{R}^2$ and $o \in \mathbf{R}^6$, such that there is z', o', w_k and w_l with $w_k = (z, o')$ and $w_l = (z', o)$. Write X_z for the projection of X onto $\mathbf{R}^2$, and X_o for the projection of X onto $\mathbf{R}^6$; so $X = X_z \times X_o$.

The following tables contain the results of calculating $r_k \cdot w_t$ (so that $r_k \cdot w_t$ is the content of the cell with row r_k and column w_t).

	w_1	w_2	w_3
r_1	$1 + M$	$\rho(\hat{z}_i^2) \cdot \hat{z}_j^1 + \varepsilon\rho(\hat{z}_i^2) \cdot e_{12}$	$\rho(\hat{z}_i^2) \cdot \hat{z}_j^1 + M$
r_2	$\rho(\hat{z}_j^1) \cdot \hat{z}_i^2$	$1 + \varepsilon\rho(\hat{z}_j^1) \cdot e_{12}$	1
r_4	$\rho(z_j^2) \cdot \hat{z}_i^2 + M$	$\rho(z_j^2) \cdot \hat{z}_j^1 + \varepsilon\rho(z_j^2) \cdot e_{12} + 2M$	$\rho(z_j^2) \cdot \hat{z}_j^1 + M$
r_5	$\rho(\hat{z}_h^1) \cdot \hat{z}_i^2$	$\rho(\hat{z}_h^1) \cdot \hat{z}_j^1 + \varepsilon\rho(\hat{z}_h^1) \cdot e_{12}$	$\rho(\hat{z}_h^1) \cdot \hat{z}_j^1$
r_7	$\rho(\hat{z}_h^2) \cdot \hat{z}_i^2 + M$	$\rho(\hat{z}_h^2) \cdot \hat{z}_j^1 + \varepsilon\rho(\hat{z}_h^2) \cdot e_{12} + 2M$	$\rho(\hat{z}_h^2) \cdot \hat{z}_j^1 + M$
r_8	$\rho(\hat{z}_i^1) \cdot \hat{z}_i^2$	$\rho(\hat{z}_i^1) \cdot \hat{z}_j^1 + \varepsilon\rho(\hat{z}_i^1) \cdot e_{12}$	$\rho(\hat{z}_i^1) \cdot \hat{z}_j^1$

	w_4	w_5
r_1	$\rho(\hat{z}_i^2) \cdot \hat{z}_j^2 + M$	$\rho(\hat{z}_i^2) \cdot \hat{z}_h^1 + \varepsilon\rho(\hat{z}_i^2) \cdot e_{12} + 2M$
r_2	$\rho(\hat{z}_j^1) \cdot \hat{z}_j^2$	$\rho(\hat{z}_j^1) \cdot \hat{z}_h^1 + \varepsilon\rho(\hat{z}_j^1) \cdot e_{12}$
r_4	$1 + M$	$\rho(z_j^2) \cdot \hat{z}_h^1 + \varepsilon\rho(z_j^2) \cdot e_{12}$
r_5	$\rho(\hat{z}_h^1) \cdot \hat{z}_j^2$	$1 + \varepsilon\rho(\hat{z}_h^1) \cdot e_{12}$
r_7	$\rho(\hat{z}_h^2) \cdot \hat{z}_j^2 + M$	$\rho(\hat{z}_h^2) \cdot \hat{z}_h^1 + \varepsilon\rho(\hat{z}_h^2) \cdot e_{12} + 2M$
r_8	$\rho(\hat{z}_i^1) \cdot \hat{z}_j^2$	$\rho(\hat{z}_i^1) \cdot \hat{z}_h^1 + \varepsilon\rho(\hat{z}_i^1) \cdot e_{12}$

	w_6	w_7	w_8	w_9
r_1	$\rho(\hat{z}_i^2) \cdot \hat{z}_h^1 + M$	$\rho(\hat{z}_i^2) \cdot \hat{z}_h^2 + M$	$\rho(\hat{z}_i^2) \cdot \hat{z}_i^1 + \varepsilon\rho(\hat{z}_i^2) \cdot e_{12} + 2M$	$\rho(\hat{z}_i^2) \cdot \hat{z}_i^1 + M$
r_2	$\rho(\hat{z}_j^1) \cdot \hat{z}_h^1$	$\rho(\hat{z}_j^1) \cdot \hat{z}_h^2$	$\rho(\hat{z}_j^1) \cdot \hat{z}_i^1 + \varepsilon\rho(\hat{z}_j^1) \cdot e_{12}$	$\rho(\hat{z}_j^1) \cdot \hat{z}_i^1$
r_4	$\rho(z_j^2) \cdot \hat{z}_h^1 + M$	$\rho(z_j^2) \cdot \hat{z}_h^2 + M$	$\rho(z_j^2) \cdot \hat{z}_i^1 + \varepsilon\rho(z_j^2) \cdot e_{12} + 2M$	$\rho(z_j^2) \cdot \hat{z}_i^1 + M$
r_5	1	$\rho(\hat{z}_h^1) \cdot \hat{z}_h^2$	$\rho(\hat{z}_h^1) \cdot \hat{z}_i^1 + \varepsilon\rho(\hat{z}_h^1) \cdot e_{12}$	$\rho(\hat{z}_h^1) \cdot \hat{z}_i^1$
r_7	$\rho(\hat{z}_h^2) \cdot \hat{z}_h^1 + M$	$1 + M$	$\rho(\hat{z}_h^2) \cdot \hat{z}_i^1 + \varepsilon\rho(\hat{z}_h^2) \cdot e_{12}$	$\rho(\hat{z}_h^2) \cdot \hat{z}_i^1 + M$
r_8	$\rho(\hat{z}_i^1) \cdot \hat{z}_h^1$	$\rho(\hat{z}_i^1) \cdot \hat{z}_h^2$	$1 + \varepsilon\rho(\hat{z}_i^1) \cdot e_{12}$	1

Define the graph (X, R) by letting $w \, R \, w'$ if and only if $w = w_k$ for some k and $r_k \cdot w_k > r_k \cdot w'$. Careful (if tedious) inspection of the calculations above, (and using inequalities (1) and (2)) reveal that

$$R = \{(w_1, w_2), (w_2, w_3), (w_4, w_5), (w_5, w_6), (w_7, w_8), (w_8, w_9)\}. \tag{3}$$

Let τ be a truth table for which our clause $C = y_i \vee y_j \vee y_h$ is true. That is: $\tau(y) = 1$ for at least one $y \in \{y_i, y_j, y_h\}$. Let B be the binary relation induced by τ (see the definition above).

Note that B is an acyclic binary relation on $\{\hat{z}_k^q : q = 1, 2, k = i, j, h\}$. None of the vectors in $\{\hat{z}_k^q : q = 1, 2, k = i, j, h\}$ is larger than the other (in the usual order on $\mathbf{R}^2$). There is therefore a function $u : \mathbf{R}_+^2 \to \mathbf{R}$ for which $u(z) > u(z')$ whenever $z \, B \, z'$ or $z > z'$.

Let R' be defined by: $w R' w'$ if (1) $w R w'$ or (2) if there are $z, z' \in X_Z$ and o such that $w = (z, o)$, $w' = (z', o)$ and $u(z) > u(z')$. Say that pair (w, w') is a 1-edge if $w \, R' \, w'$ for reason (1) and a 2-edge if $w \, R' \, w'$ for reason (2).

Lemma 2. *(X, R') is acyclic.*

Proof. Suppose, towards a contradiction, that there is a cycle. The cycle cannot consist purely of 2-edges because each 2-edge implies an increase in $u(z)$. Therefore some of the edges must be 1-edges. Inspection of the graph (X, R) reveals that all of the edges in R must then be part of this cycle. The reason is that edges can connect (z, o) and (z', o') with $o \neq o'$ only if they belong to R.[3] Then a cycle can only be closed if it involves *all* of the edges in R. Such a cycle would define a path from w_9 to w_1, from w_3 to w_4, and from w_6 to w_7. Each of these paths would involve only 2-edges. By definition of 2-edges, then $u(w_1) > u(w_9)$, $u(w_4) > u(w_3)$, and $u(w_7) > u(w_6)$. But $u(w_1) > u(w_9)$ can only be true if $\hat{z}_i^2 \, B \, \hat{z}_i^1$. Similarly, we obtain that $\hat{z}_j^2 \, B \, \hat{z}_j^1$ and $\hat{z}_h^2 \, B \, \hat{z}_h^1$. This contradicts that C is true under the truthtable τ. $\square$

[3]While the extension of B defined by u is arbitrary, this does not cause new edges unless two bundles share their o component. For example, while it may be the case that $u(\hat{z}_h^2) > u(\hat{z}_j^2)$, this does not mean that $w_7 \, R' \, w_4$ as w_7 and w_4 differ in their third and eight components.

Lemma 3. *There is a function v such that $v(u(z), o) > v(u(z'), o')$ whenever $(z, o)\ R'\ (z', o')$ or $u(z) \geq u(z')$ and $o > o'$.*

Proof. Let R'' be a binary relation on $u(\mathbf{R}^2_+) \times \mathbf{R}^4_+$ defined by $(s, o)\ R''\ (s', o')$ if $(s, o) > (s', o')$ or if there is z and z' such that $s = u(z)$ and $s' = u(z')$ and $(z, o)\ R\ (z, o)$. By Lemma 2 and the observation that non of the vectors in X_o is comparable in the usual Euclidean order, the relation R'' is acyclic. The set $u(X_z) \times X_o \cup \mathbb{Q} \cap u(\mathbf{R}^2_+) \times \mathbb{Q}^4_+$ is countable and order dense, so there is a function v as required by the statement of the lemma. □

3.2.2. *Step 2: The construction for L clauses.* For each clause C_l, define the bundles w_t and prices r_k as above. Let w^l_t be defined as the vector in $\mathbf{R}^9$ obtained by embedding w_t and adding $M2^l e_9$. Let r^l_k be the sum of e_9 and the embedding of r_k in $\mathbf{R}^9$. As a result of these definitions, $r^l_k \cdot w^k_t = r_k \cdot w_t + M2^l$. Importantly, for a fixed l, the comparison of $r^l_k \cdot w^k_k$ and $r^l_k \cdot w^k_t$ is the same as the comparison of $r_k \cdot w_k$ and $r_k \cdot w_t$ performed in Step 1.

Define $X^l = X^l_z \times X^l_o$ from $w^l_1, \ldots, w^l_9$ in the same way as $X = X_z \times X_o$ was defined. Define R^l from w^l_t and r^l_k in the same way as R was defined. Note that $r^l_k \cdot w^k_t$ only differ from $r_k \cdot w_t$ in the constant $M2^l$, so the graphs (X, R) and (X^l, R^l) are the same once we identify w_k with w^l_k.

Let $\bar{X} = \cup^L_{l=1} X^l$. Define a binary relation $\bar{R}$ on $\bar{X}$ by $w\,\bar{R}\,w'$ iff there is some l and some t such that $w = w^l_t$ and $r^l_t \cdot w^l_t > r^l_t \cdot w'$. Note that $\bar{R}$ has the following properties

(1) $\bar{R}$ coincides with R^l on X^l ($R^l = \bar{R} \cap X^l \times X^l$);
(2) if $w \in X^l$ and $w' \in X^{l'}$, with $l' < l$, then $w\,\bar{R}\,w'$ and it is false that $w'\,\bar{R}\,w$.

Let τ be a truthtable for which all clauses C^l are true. Such a truth table defines a binary relation B on $\bar{X}_z = \cup_l X^l_z$. The binary relation is acyclic, as there are no pairs of subsequent edges in B. As in Step 1, there is a function $u : \mathbf{R}^2_+ \to \mathbf{R}$ for which $u(z) > u(z')$ whenever $z\,B\,z'$ or $z > z'$.

Let $\bar{R}'$ be defined by: $w\,R'\,w'$ if (1) $w\,R\,w'$ or (2) if there are $z, z' \in \bar{X}_z$ and o such that $w = (z, o)$, $w' = (z', o)$ and $u(z) > u(z')$.

Lemma 4. *$(\bar{X}, \bar{R}')$ is acyclic.*

Proof. By the second property of $\bar{R}$, there cannot exist a cycle that contains an edge going from $w \in X^l$ to $w' \in X^{l'}$ with $l \neq l'$. Therefore any cycle must contain only vertexes in some X^l. By Lemma 1, there is no such cycle. □

The following result, which finishes the proof, follows from Lemma 4 in a similar way to how Lemma 3 follows from Lemma 1.

Lemma 5. *There is a function* v *such that* $v(u(z), o) > v(u(z'), o')$ *whenever* $(z, o)\ \bar{R}'\ (z', o')$ *or* $u(z) \geq u(z')$ *and* $o > o'$.

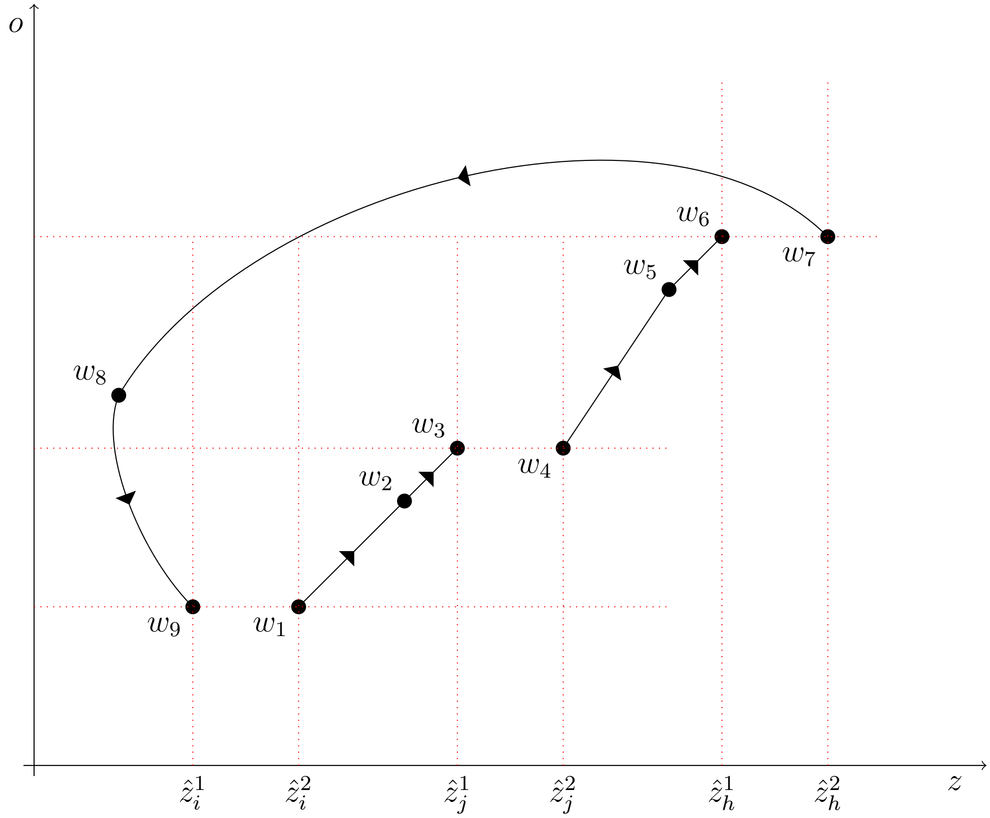

Figure 1. The construction for a single clause.

3.3. **Some remarks on the proof.** 1) The construction for one clause can be summarized in the diagram depicted in Figure 1. The horizontal axis represents $\mathbf{R}^2$ and the vertical axis $\mathbf{R}^6$. The directions of the arrows reflect the binary relation R: $w_1\, R\, w_2$ is denoted by the arrow pointing from w_1 to w_2, and so on. Note that the pairs of bundles w_9 and w_1, w_3 and w_4, and w_6 and w_7 share their o component. These pairs are the only ones that share an o component.

2) The presence of the bundles w_2, w_5 and w_8 may need an explanation. We need to use them for the following reason. Consider the case of w_2. We want to have $w_1\, R\, w_3$, but not that $w_1\, R\, w_4$. This is difficult

because w_3 and w_4 differ only in the z component. By introducing w_2, which dominates w_3 but not w_4, we can achieve the desired relations.

3) I have taken a shortcut in the proof by introducing exponential quantities w_t^l. They are there to make sure that certain quantities are large enough, and are easily avoided.

4) The definition of the bundles z_k^q and supporting prices $\rho(z_i^q)$ may involve using irrational numbers, which is questionable from al algorithmic viewpoint. Since the inequalities in Lemma 1 are strict, these numbers can be replaced with rational numbers to have the construction only operate with "discrete" objects. (In a similar fashion, the primitive dataset should only involve consumption bundles and prices with rational entries.)

5) The main contribution here is to do the construction for a fixed number of goods. If one is free to use any number of different goods to capture the different edges, then it is easy to recreate any given graph as a revealed preference binary relation. When the number of goods is fixed, not all graphs can be revealed preference relations (the best known example is the case when there are two goods, in which the weak axiom of revealed preference suffices for rationalizability). So it is important to be able to work with a rather specific graph, the one depicted in Figure 1. The ability to do the reduction for a fixed number of goods relies, among other things, on using 3SAT.

References

Cherchye, L., T. Demuynck, and B. De Rock (2011): "Revealed preference tests for weak separability: an integer programming approach," Forthcoming in the *Journal of Econometrics*.

Deaton, A. S. (1974): "The analysis of consumer demand in the United Kingdom, 1900-1970," *Econometrica*, pp. 341–367.

Quah, J. (2012): "A revealed preference test for weakly separable preferences," Mimeo, Department of Economics, University of Oxford.

Varian, H. R. (1983): "Non-parametric Tests of Consumer Behaviour," *The Review of Economic Studies*, 50(1), 99–110.