# Roberts' Theorem with Neutrality: A Social Welfare Ordering Approach *

Debasis Mishra[†] and Arunava Sen [‡]

November 23, 2011

**Abstract**

We consider dominant strategy implementation in private values settings, when agents have multi-dimensional types, the set of alternatives is finite, monetary transfers are allowed, and agents have quasi-linear utilities. We focus on private-value environments. We show that any implementable and neutral social choice function must be a weighted welfare maximizer if the type space of every agent is an $m$-dimensional open interval, where $m$ is the number of alternatives. When the type space of every agent is unrestricted, Roberts' Theorem with neutrality (Roberts, 1979) becomes a corollary to our result. Our proof technique uses a *social welfare ordering* approach, commonly used in aggregation literature in social choice theory. We also prove the general (affine maximizer) version of Roberts' Theorem for unrestricted type spaces of agents using this approach.

Keywords: Dominant strategy mechanism design; Roberts' Theorem; affine maximizers; social welfare ordering

JEL Classification: D44.

*This paper has benefited from detailed comments of two anonymous referees and an associate editor. The authors are extremely grateful to Sushil Bikhchandani, Juan Carlos Carbajal, Claude d'Aspremont, Mridu Prabal Goswami, Amit Goyal, Ron Lavi, Thierry Marchant, Hervé Moulin, Arup Pal, David Parkes, Kevin Roberts, Maneesh Thakur, Rakesh Vohra, and seminar participants at the "Multidimensional Mechanism Design" workshop in Bonn, Universitat Autònoma de Barcelona, Delhi School of Economics, Social Choice and Welfare Meeting, and Econometric Society World Congress for their comments.

†Corresponding Author. Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi - 110016.

‡Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi - 110016.

# 1 Introduction

This paper is a contribution to the literature that investigates the structure of social choice functions which can be implemented in dominant strategies in settings where monetary transfers are allowed and the underlying utility function of every agent is quasi-linear in money. Vickrey (1961); Clarke (1971); Groves (1973) showed that efficient social choice functions can be implemented by a unique family of transfer rules, now popularly known as Vickrey-Clarke-Groves (VCG) transfer schemes. Remarkably, when the domain is unrestricted (as in the Gibbard-Satterthwaite setup) and the range of the mechanism contains at least three alternatives, the only (dominant strategy) implementable social choice functions are *affine maximizers*. These social choice functions are generalizations of weighted efficiency rules. This result was proved by Roberts (1979) in a seminal paper. It can be seen as the counterpart to the Gibbard-Satterthwaite theorem for quasi-linear utility environments.

As in the literature without money, the literature with quasi-linear utility has since tried to relax various assumptions in Roberts' Theorem. Rochet (1987) shows that a certain *cycle monotonicity* property characterizes dominant strategy implementable social choice functions. Though this characterization is very general - works for any domains and any set of alternatives (finite or infinite) - it is not as useful as the Roberts' Theorem since it does not give a functional form of the class of implementable social choice functions. Along the lines of Rochet (1987), Bikhchandani et al. (2006) and Saks and Yu (2005) have shown that a *weak monotonicity* property characterizes implementable social choice functions in auction settings, a severely restricted domain, when the set of alternatives is finite and the type space is convex [1]. Again, the precise functional form of the implementable social choice functions are missing in these characterizations [2]. A fundamental open question is the following:

*What subdomains imply that an implementable social choice function is an affine maximizer or a variant?*

Part of the reason this question is not completely answered is that Roberts' original proof is somewhat mysterious. Several attempts have been made recently to simplify, refine, and extend Roberts' Theorem. Using almost the same structure and approach, Lavi et al. (2009) reduced the complexity of Roberts' original proof. Dobzinski and Nisan (2009) also provide an alternate (modular) proof of Roberts' Theorem for unrestricted domain. Building on the technique of Lavi et al. (2009), Carbajal et al. (2011) extend Roberts' theorem to a general domain which they term "comprehensive domain". Other proofs of Roberts' Theorem can

[1] See also Ashlagi et al. (2010) who prove the converse of this result, and Gui et al. (2004) who introduced a graph-theoretic foundation of incentive compatibility constraints. Muller et al. (2007) provide a monotonicity characterization for Bayes-Nash incentive compatibility.

[2] Parallel monotonicity characterizations have been done when the set of alternatives is infinite in Archer and Kleinberg (2008), and for general type (valuation) functions by Berger et al. (2010). An excellent survey of these results is found in Vohra (2011).

be found in (for unrestricted domains) Jehiel et al. (2008) and Vohra (2008).

To our knowledge, there has been no "functional-form" characterization of implementable social choice functions in restricted domains. One particular characterization of Lavi et al. (2003) stands out. They focus on a particular restricted domain, which they call *order-based domains*. Under various additional restrictions on the social choice function, they show that every implementable social choice function must be an "almost" affine maximizer - roughly, almost affine maximizers are affine maximizers for large enough values of types of agents.

## 1.1 Our Contribution

Our paper contributes to the literature in two ways. First, we characterize restricted domains where the affine maximizer theorem holds in the presence of an additional assumption on social choice functions, that of *neutrality*. Neutrality requires the social choice function to treat all alternatives symmetrically. It is a familiar axiom in social choice theory and we discuss it at greater length in Section 3. Our main result states that every implementable social choice function is a weighted welfare maximizer if the type space of every agent is an $m$-dimensional open interval, where $m$ is the number of alternatives. For the unrestricted domain, our result implies Roberts' result in the special case where attention is restricted to neutral social choice functions. We demonstrate that the neutrality assumption is essential for our domain characterization result in the following sense: there exist (open interval) domains over which an implementable non-affine-maximizer social choice function exists but over which all neutral implementable social choice functions are weighted welfare maximizers. None of the existing results in the literature imply our result.

Our second contribution is methodological and conceptual. Our proof technique differs significantly from existing ones. It can be summarized in three steps.

- S1 We show that an implementable and neutral social choice function induces an ordering on the domain [3].
- S2 This ordering satisfies three key properties: *weak Pareto, translation invariance, and continuity*.
- S3 We then prove a result on the representation of any ordering which satisfies these properties. For unrestricted domains this result is familiar in the literature - see for instance, Blackwell and Grishick (1954), d'Aspremont and Gevers (1977), Blackorby et al. (1984), Trockel (1992) and d'Aspremont and Gevers (2002). We show that any ordering on an open and convex set which satisfies the axioms specified in S2 can be represented by a weighted welfare maximizer.

[3] Recall that an ordering on a set $X$ is a reflexive, complete, and transitive binary relation of the elements of $X$.

The key feature of our approach is to transform the problem of characterizing incentive-compatible social choice functions over a domain into a particular problem of characterizing orderings of vectors in that domain. The problem of characterizing orderings satisfying properties such as weak Pareto, invariance, continuity etc (over the unrestricted domain), is a classical one in social choice theory. It arose from the recognition that a natural way to escape the negative conclusions of the Arrow Impossibility Theorem is to enrich the informational basis of Arrovian social welfare functions from individual preference orderings to utility functions. Constructing a social welfare function that satisfies the (standard) axioms of Independence of Irrelevant Alternatives and Pareto Indifference, is equivalent to constructing an ordering over $\mathbb{R}^n$ where $n$ is the number of individuals. The aggregation problem in this environment can therefore be reduced to the problem of determining an appropriate ordering of the vectors in $\mathbb{R}^n$. There is an extensive literature that investigates precisely this question (see d'Aspremont and Gevers (2002) for a comprehensive survey).

It has been known that there is a deep connection between two seemingly unrelated problems in social choice - the strategic problem with the goal of characterizing incentive-compatible social choice functions and the aggregation problem with the objective of characterizing social welfare functions satisfying the Arrovian axioms. For instance in the case of the unrestricted domain consisting of all preference orderings, the Arrow Theorem can be used to prove the Gibbard-Satterthwaite Theorem and vice-versa (for a unified approach to both problems see Reny (2001)). Our proof serves to highlight this connection further by demonstrating the equivalence of a strategic problem in a quasi-linear domain with an aggregation problem involving utility functions.

We also remark that though the representation result in Step S3 is well-known for unrestricted domains, our extension to open and convex domain may be of some independent interest.

Finally, we show how Roberts' affine maximizer theorem (for unrestricted domain) can be proved using Roberts' Theorem with neutrality. We leverage our result with neutrality to prove this. This proof is contained in Section 7.

## 2 Roberts' Affine Maximizer Theorem

Let $A = \{a, b, c, \ldots\}$ be a finite set of alternatives or allocations. Suppose $|A| = m \geq 3$. Let $N = \{1, \ldots, n\}$ be a finite set of agents. The type of agent $i$ is a vector in $\mathbb{R}^m$. Denote by $t_i$ the type (vector) of agent $i \in N$, where for every $a \in A$, $t_i^a$ denotes the *value* of agent $i$ for alternative $a$ when his type is $t_i$ - this is the standard *private values* model, where every agent knows his values for alternatives. A type profile will be denoted by $t$, and consists of $n$ vectors in $\mathbb{R}^m$. Alternately, one can view a type profile $t$ to be an $n \times m$ matrix, where every row represents a type vector of an agent. The column vectors are vectors in $\mathbb{R}^n$. We refer to a column vector generated by a type profile to be a utility vector. Hence, $t^a$ represents the utility vector corresponding to allocation $a$ in type profile $t$ and $t^{-a}$ will denote the utility vectors in type profile $t$ except $t^a$.

Let $T_i$ be the type space (the set of all type vectors) of agent $i$. We assume $T_i = (\alpha_i, \beta_i)^m$ where $\alpha_i \in \mathbb{R} \cup \{-\infty\}$, $\beta_i \in \mathbb{R} \cup \{\infty\}$, and $\alpha_i < \beta_i$. We call such a type space an **$m$-dimensional open interval domain**. The set of all type profiles is denoted by $T = T_1 \times T_2 \times \ldots T_n$.

Let the set of all utility vectors for every alternative in $A$ be $D \subseteq \mathbb{R}^n$, which is an open rectangle in $\mathbb{R}^n$. Hence, the set of type profiles can be alternatively written as $D^m$. Throughout, we will require different mathematical properties of $D$ which are satisfied by $T_i$ for every $i$ if it is an $m$-dimensional open interval domain. In particular, note the following two properties which hold if the type space is an $m$-dimensional open interval:

1. If we have a type profile $t$ in our domain and permute two utility vectors $t^a$ and $t^b$ in this type profile, then we will get a valid type profile in our domain.

2. For every type profile $t$ in our domain and every $a \in A$, there exists $\epsilon \gg 0$ [4] such that if we increase the utility vector $t^a$ by $\epsilon$, then we get a valid type profile in our domain.

The first property follows from the interval assumption and the second property follows from the openness assumption. We use these two properties extensively in our proofs.

We use the standard notation of $t_{-i}$ to denote a type profile of agents in $N \setminus \{i\}$ and $T_{-i}$ to denote the type spaces of agents in $N \setminus \{i\}$.

A social choice function is a mapping $f : T \to A$. A payment function is a mapping $p : T \to \mathbb{R}^n$. The payment of agent $i$ at type profile $t$ is denoted by $p_i(t)$.

Definition 1 *A social choice function $f$ is implementable (in dominant stragies) if there exists a payment function $p$ such that for every $i \in N$ and every $t_{-i}$, we have*

$$t_i^{f(t_i,t_{-i})} + p_i(t_i, t_{-i}) \geq t_i^{f(s_i,t_{-i})} + p_i(s_i, t_{-i}) \qquad \forall\ s_i, t_i \in T_i.$$

*In this case, we say that $p$ implements $f$.*

Every social choice function satisfies certain properties if it is implementable. Below, we give one such useful property.

Definition 2 *A social choice function $f$ satisfies* **positive association of differences (PAD)** *if for every $s, t \in T$ such that $f(t) = a$ with $s^a - t^a \gg s^b - t^b$ for all $b \neq a$, we have $f(s) = a$.*

Lemma 1 (**Roberts** (**1979**)) *Every implementable social choice function satisfies PAD.*

[4] For every pair of vectors $x, y \in \mathbb{R}^n$, we say that $x \gg y$ if and only if $x$ is greater than $y$ in every component.

A natural question to ask is what social choice functions are implementable. In an important result, Roberts (1979) characterized the set of all social choice functions when the type space is unrestricted and when the social choice function satisfies a condition called *non-imposition*.

Definition 3 *A social choice function $f$ satisfies* **non-imposition** *if for every $a \in A$, there exists $t \in T$ such that $f(t) = a$.*

Using PAD and non-imposition, Roberts proved the following theorem.

Theorem 1 (**Roberts (1979)**) *Suppose $T_i = \mathbb{R}^m$ for all $i \in N$. If $f$ is an implementable social choice function and satisfies non-imposition, then there exists weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ and a function $\kappa : A \to \mathbb{R}$ such that for all $t \in T$,*

$$f(t) \in \arg\max_{a \in A} \Big[ \sum_{i \in N} \lambda_i t_i^a - \kappa(a) \Big]$$

The social choice functions in this family are called **affine maximizer social choice functions**. Below, we give an example with two agents to illustrate that Theorem 1 does not hold in bounded domains. In the example below, it is essential to assume that there are at least two agents. Roberts' Theorem holds in any bounded domain if there is only a single agent (Chung and Ely, 2006). Chung and Ely (2006) refer to this characterization as *pseudo-efficiency*.

### 2.1 Non-Affine-Maximizers in Bounded Domains: An Example

Here, we give an example to illustrate that Theorem 1 does not hold in bounded domains. The example is due to Meyer-ter-Vehn and Moldovanu (2003).

Example 1

Let $N = \{1, 2\}$ and $A = \{a, b, c\}$. Suppose $T_1 = T_2 = (0, 1)^3$ (alternatively, suppose $D = (0, 1)^2$). Consider the following allocation rule $f$. Let

$$\mathbb{T}^g = \{(t_1, t_2) \in \mathbb{T}^2 : t_1^c < t_1^b + 0.5\} \cup \{(t_1, t_2) \in \mathbb{T}^2 : t_2^c > t_2^b - 0.5\}.$$

Then,

$$f(t_1, t_2) = \begin{cases} \arg\max\{-1.5 + t_1^a + t_2^a, t_1^b + t_2^b, t_1^c + t_2^c\} & \forall\ (t_1, t_2) \in \mathbb{T}^g \\ c & \forall\ (t_1, t_2) \in \mathbb{T}^2 \setminus \mathbb{T}^g. \end{cases}$$

It can be verified that $f$ satisfies non-imposition. Further, the following payment rule $p$ implements $f$.

$$p_1(t_1, t_2) = \begin{cases} t_2^a & \text{if } f(t_1, t_2) = a \\ \min\{1.5 + t_2^b, 2 + t_2^c\} & \text{if } f(t_1, t_2) = b \\ (1.5 + t_2^c) & \text{if } f(t_1, t_2) = c. \end{cases}$$

$$p_2(t_1,t_2) = \begin{cases} t_1^a & \text{if } f(t_1,t_2)=a \\ (1.5+t_1^b) & \text{if } f(t_1,t_2)=b \\ \min\{1.5+t_1^c, 2+t_1^b\} & \text{if } f(t_1,t_2)=c. \end{cases}$$

However, one can verify that $f$ is not an affine maximizer.

## 3 The Main Result

In this section, we state our main result, and prove it in subsequent sections. We restrict attention to neutral social choice functions which we now describe. Neutrality roughly requires that the mechanism designer should treat all allocations in $A$ symmetrically. The standard notion of neutrality, which we call *scf-neutrality* is the following.

Definition 4 *A social choice function $f$ is* **scf-neutral** *if for every $t \in T$, every permutation $\rho$ of $A$ and type profile $s$ induced by permutation $\rho$ on $t$, we have $f(s) = \rho(f(t))$ if $t \neq s$.*

The scf-neutrality requires the social choice function not to discriminate between alternatives. It is extensively used in the social choice theory literature (Moulin, 1983). Consider a city planner who has to choose amongst several projects which are ex-ante symmetric. For instance, if the decision is regarding selecting a firm which will execute a public project, then discriminating between firms may be prohibited by law. Also, note that neutrality is a natural assumption in voting models, where the mechanism designer will not like to discriminate among candidates. Indeed, Roberts' (Roberts, 1979) initial motivation was to examine implementability by introducing monetary transfers in voting models.

We require a slightly different version of neutrality, which we find easier to use. For this, we require the notion of a choice set, which we define below. Given a social choice function $f$ we define the following set. For every $t \in T$, the **choice set** at $t$ is defined as:

$$C^f(t) = \{a \in A: \ \forall\, \epsilon \gg 0 \text{ such that } (t^a+\epsilon, t^{-a}) \in T, f(t^a+\epsilon, t^{-a}) = a\}.$$

We first show that choice sets are non-empty under our assumptions of the domain ($m$-dimensional open intervals).

Lemma 2 *Let $f$ be an implementable social choice function. Then, for every $t \in T$, $f(t) \in C^f(t)$.*

*Proof*: Consider $t \in T$, and let $f(t) = a$. Choose $\epsilon \gg 0$ such that $(t^a+\epsilon, t^{-a}) \in T$ (since $D$ is open, this is possible). By PAD, $f(t^a+\epsilon, t^{-a}) = a$. Hence, $a \in C^f(t)$. ■

Using the notion of a choice set, we define a neutral social choice function.

Definition 5 *A social choice function $f$ is* **neutral** *if for every type profile $t \in T$, every permutation $\rho$ of $A$ and type profile $s$ induced by permutation $\rho$* [5]*, we have $C^f(s) = \{\rho(a) : a \in C^f(t)\}$.*

This version of neutrality requires that if we permute the set of alternatives, and go from type profile $t$ to $s$, then the choice set at $s$ is just the set of alternatives obtained by applying the permutation to the choice set at $t$. This neutrality is weaker than scf-neutrality in the presence of implementability.

Claim 1 *If a social choice function $f$ is implementable and scf-neutral, then it is neutral.*

*Proof*: Since $f$ is implementable, it satisfies PAD. Fix a type profile $t$ and a permutation $\rho$ of $A$, and let $s$ be the type profile induced by $\rho$ on $t$. Consider $a \in C^f(t)$ and a type profile $u = (u^a = t^a + \epsilon, u^{-a} = t^{-a})$ for some $\epsilon \in \mathbb{R}^n_{++}$. Hence, $f(u) = a$. Now, let $v$ be the type profile induced by permutation $\rho$ on $u$. By scf-neutrality $f(v) = \rho(a)$ ($\epsilon$ can be chosen arbitrarily small so that $u \neq v$). By PAD, $\rho(a) \in C^f(s)$.

To show that for any $a \notin C^f(t)$, we must have $a \notin C^f(s)$, assume for contradiction $a \in C^f(s)$, and apply the previous argument to conclude $a \in C^f(t)$. This gives the desired contradiction. ■

Non-imposition is implied by neutrality in $m$-dimensional open interval domains.

Lemma 3 *Suppose $f$ is an implementable social choice function. If $f$ is neutral then it satisfies non-imposition.*

*Proof*: Fix an alternative $a \in A$. Consider any arbitrary type profile $t$ such that $f(t) = b \neq a$. By Lemma 2, $b \in C^f(t)$. Now, construct another type profile $s = (s^a = t^b, s^b = t^a, s^{-ab} = t^{-ab})$. By neutrality, $a \in C^f(s)$. Now, let $u = (u^a = s^a + \epsilon, u^{-a} = s^{-a})$ for some $\epsilon \in \mathbb{R}^n_{++}$. Since $a \in C^f(s)$, we have that $f(u) = a$. Hence, $f$ satisfies non-imposition. ■

Under neutrality, Roberts' Theorem is modified straightforwardly as follows (see Lavi (2007)).

Theorem 2 (**Roberts (1979)**) *Suppose $T_i = \mathbb{R}^m$ for all $i \in N$. If $f$ is an implementable social choice function and satisfies neutrality, then there exists weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ such that for all $t \in T$,*

$$f(t) \in \arg\max_{a \in A} \sum_{i \in N} \lambda_i t_i^a$$

[5] Given a permutation $\rho$ of $A$ and a type profile $t$, the type profile induced by permutation $\rho$ is the profile obtained from $t$ by relabeling the columns based on permutation $\rho$.

A striking aspect of this theorem is that it gives a precise functional form of the neutral social choice functions that can be implemented. This family of social choice functions is called the **weighted welfare maximizer social choice functions**. If all the weights ($\lambda_i$s) are equal in a weighted welfare maximizer social choice function, then we get the **efficient social choice function**.

We are now ready to state our main result. It says that if the type space of every agent is an $m$-dimensional open interval, then the every neutral and implementable social choice function is a weighted welfare maximizer.

THEOREM **3** *Suppose $f$ is a neutral social choice function and for every $i \in N$, $T_i$ is an $m$-dimensional open interval. If the social choice function $f$ is implementable, then there exists weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ such that for all $t \in T$,*

$$f(t) \in \arg\max_{a \in A} \sum_{i \in N} \lambda_i t_i^a.$$

The converse of Theorem 3 is not true. Carbajal et al. (2011) give an example which illustrates that not every weighted welfare maximizer can be implemented. However, if we impose a mild condition on a weighted welfare maximizer then it can be implemented.

DEFINITION **6** *A weighted welfare maximizer $f$ with weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ is* **unresponsive to irrelevant agents (UIA)** *if for every $i \in N$ with $\lambda_i = 0$, we have for every $t_{-i}$ and for every $s_i, t_i \in T_i$, $f(s_i, t_{-i}) = f(t_i, t_{-i})$.*

Now, one can verify that if $f$ is a weighted welfare maximizer with weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ and is UIA, then the following payment function $p : T \to \mathbb{R}^n$ makes the social choice function implementable. For all $i \in N$ with $\lambda_i = 0$, $p_i(t) = 0$ for all $t \in T$. For all $i \in N$ with $\lambda_i > 0$,

$$p_i(t) = \frac{1}{\lambda_i}\Big[\sum_{j \neq i} \lambda_j t_j^{f(t)}\Big] - h_i(t_{-i}) \qquad \forall\, t \in T.$$

where $h_i : \mathbb{T}_{-i} \to \mathbb{R}$.

In Sections 4, 5, and 6, we describe various constructs and results that lead to various steps in the proof of Theorem 3.

## 4 An Induced Social Welfare Ordering

In the aggregation theory literature, an axiom called "binary independence" is extensively used - see d'Aspremont and Gevers (2002). Roughly, it says that the comparision between two alternatives $a$ and $b$ should only depend on the utility (column) vectors corresponding to $a$ and $b$. We prove a counterpart of this axiom for our choice set for $m$-dimensional open interval domains.

Proposition 1 **(Binary Independence)** *Let $f$ be an implementable social choice function. Consider two type profiles $t = (t^a, t^b, t^{-ab})$, $s = (s^a = t^a, s^b = t^b, s^{-ab})$.*

*a) Suppose $a, b \in C^f(t)$. Then, $a \in C^f(s)$ if and only if $b \in C^f(s)$.*

*b) Suppose $a \in C^f(t)$ but $b \notin C^f(t)$. Then $b \notin C^f(s)$.*

*Proof*: Suppose $a, b \in C^f(t)$. Now, consider a type profile $u = (u^a = t^a, u^b = t^b, u^{-ab})$, where $u_i^c = \min(t_i^c, s_i^c)$ for all $i \in M$ and for all $c \notin \{a, b\}$. Note that since $D$ is an open rectangle in $\mathbb{R}^n$, $u \in D^m$.

a) Suppose $a, b \in C^f(t)$. We will first show that $a, b \in C^f(u)$. Choose an $\epsilon \gg 0$. Since $a \in C^f(t)$, we know that $f(t^a + \frac{\epsilon}{2}, t^b, t^{-ab}) = a$. By PAD, $f(t^a + \epsilon, t^b, u^{-ab}) = a$. Hence, $a \in C^f(u)$. Using an analogous argument, $b \in C^f(u)$.

   Now, suppose that $a \in C^f(s)$ and assume for contradiction $b \notin C^f(s)$. Choose an $\epsilon \gg 0$ and arbitrarily close to zero. We show that $f(t^a + 2\epsilon, t^b + 3\epsilon, s^{-ab}) \neq b$. Assume for contradiction, $f(t^a + 2\epsilon, t^b + 3\epsilon, s^{-ab}) = b$. By PAD, $f(t^a, t^b + 4\epsilon, s^{-ab}) = b$. Since $\epsilon$ can be made arbitrarily small, this implies that $b \in C^f(s)$. This is a contradiction.

   Next, we show that $f(t^a + 2\epsilon, t^b + 3\epsilon, s^{-ab}) \neq c$ for any $c \notin \{a, b\}$. Assume for contradiction $f(t^a + 2\epsilon, t^b + 3\epsilon, s^{-ab}) = c$ for some $c \notin \{a, b\}$. By PAD, $f(t^a + 2\epsilon, t^b, s^c + \frac{\epsilon}{2}, s^{-abc}) = c$. Also, since $a \in C^f(s)$, we know that $f(t^a + \epsilon, t^b, s^c, s^{-abc}) = a$. By PAD, $f(t^a + 2\epsilon, t^b, s^c + \frac{\epsilon}{2}, s^{-abc}) = a$. This is a contradiction.

   Hence, $f(t^a + 2\epsilon, t^b + 3\epsilon, s^{-ab}) = a$. By PAD, $f(t^a + \frac{5\epsilon}{2}, t^b + 3\epsilon, u^{-ab}) = a$. We show that $f(t^a, t^b + \epsilon', u^{-ab}) \neq b$ for all $0 \ll \epsilon' \ll \frac{\epsilon}{2}$. Assume for contradiction $f(t^a, t^b + \epsilon', u^{-ab}) = b$ for some $0 \ll \epsilon' \ll \frac{\epsilon}{2}$. By PAD, $f(t^a + \frac{5\epsilon}{2}, t^b + 3\epsilon, u^{-ab}) = b$. This is a contradiction. Hence, $f(t^a, t^b + \epsilon', u^{-ab}) \neq b$ for some $\epsilon' \gg 0$. This implies that $b \notin C^f(u)$, which is a contradiction. Hence, $a \in C^f(s)$ implies that $b \in C^f(s)$.

   Using a symmetric argument, $b \in C^f(s)$ implies $a \in C^f(s)$. Hence, if $a \notin C^f(s)$ then $b \notin C^f(s)$. This implies that either $\{a, b\} \subseteq C^f(s)$ or $\{a, b\} \cap C^f(s) = \emptyset$.

b) Suppose $a \in C^f(t)$ but $b \notin C^f(t)$. As in part (a), $a \in C^f(u)$. Now, assume for contradiction, $b \in C^f(s)$. If $a \notin C^f(s)$, then exchanging the role of $a$ and $b$ in the second half of (a), we get that $a \notin C^f(u)$. This is a contradiction. If $a \in C^f(s)$, then we have $a, b \in C^f(s)$ but $a \in C^f(t)$. By part (a), $b \in C^f(t)$. This is a contradiction.

■

We will define an ordering on $D$ induced by an implementable social choice function. In general, we will refer to an arbitrary ordering $R$ on $D$. The symmetric component of an ordering $R$ will be denoted as $I$ and the anti-symmetric component will be denoted as $P$. Note that a social choice function $f$ is a mapping $f : T \to A$. Hence, for every type profile $t$, a social choice function can be thought of as picking a column vector (which belongs to $D$) in $t$. We will show that in the process of picking these column vectors in $D$ in an "implementable manner", a neutral social choice function induces a social welfare ordering.

The following is a useful lemma that we will use in the proofs.

Lemma 4 *Suppose $f$ is an implementable and neutral social choice function. Consider a type profile $t \in T$ such that $t^a = t^b$ for some $a, b \in A$. Then, $a \in C^f(t)$ if and only if $b \in C^f(t)$.*

*Proof*: This follows from the fact that permuting columns $a$ and $b$ in $t$ produces $t$ again. Hence, by neutrality, $a \in C^f(t)$ if and only if $b \in C^f(t)$. ■

Definition 7 *A* **social welfare ordering** *$R^f$ induced by a social choice function $f$ is a relation on $D$ defined as follows. The symmetric component of $R^f$ is denoted by $I^f$ and the antisymmetric component of $R^f$ is denoted by $P^f$. Pick $x, y \in D$.*

*We say $xP^f y$ if and only if there exists a profile $t$ with $t^a = x$ and $t^b = y$ for some $a, b \in A$ such that $a \in C^f(t)$ but $b \notin C^f(t)$.*

*We say $xI^f y$ if and only if there exists a profile $t$ with $t^a = x$ and $t^b = y$ for some $a, b \in A$ such that $a, b \in C^f(t)$.*

Proposition 2 **(Social Welfare Ordering)** *Suppose $f$ is an implementable and neutral social choice function. Then, the relation $R^f$ induced by $f$ on $D$ is an ordering.*

*Proof*: We first show that $R^f$ is well-defined. Pick $x, y \in D$. We consider two cases.

Case 1: Suppose $xP^f y$. Then there exists a type profile $t$ and some $a, b \in A$ with $t^a = x$ and $t^b = y$ such that $a \in C^f(t)$ but $b \notin C^f(t)$. For $R^f$ to be well-defined, we need to show that for *any* $c, d \in A$ and any $s \in T$ such that $s^c = x, s^d = y$, we have $d \notin C^f(s)$.

Consider any other type profile $s$ such that $s^a = x$ and $s^b = y$. By Proposition 1, $b \notin C^f(s)$. Consider any other profile $u$ and $(c, d) \neq (a, b)$ such that $u^c = x$ and $u^d = y$. We can permute $u$ to get another profile $v$ such that $v^a = x$ and $v^b = y$. By Proposition 1, $b \notin C^f(v)$. By neutrality, $d \notin C^f(u)$.

Case 2: Suppose $xI^f y$. Then there exists a type profile $t$ and some $a, b \in A$ such that $a, b \in C^f(t)$. Consider any other type profile $s$ such that $s^a = x$ and $s^b = y$. By Proposition 1, $a \in C^f(s)$ if and only if $b \in C^f(s)$. By neutrality (as in Case 1), the choice of $a$ and $b$ is without loss of generality. This shows that $I^f$ is well-defined.

We next show that $R^f$ is reflexive. Consider $x \in D$ and the profile where $t^a = x$ for all $a \in A$. By Lemma 4, $C^f(t) = A$. Hence, $xI^f x$.

Next, we show that $R^f$ is complete. Choose $x, y \in D$. Consider a type profile $t$ where each column vector is either $x$ or $y$ with at least one column vector being $x$ and at least one column vector being $y$. Suppose $f(t) = a$. Then either $t^a = x$ or $t^a = y$. Without loss of generality, let $t^a = x$. Take any $b \neq a$ such that $t^b = y$. If $b \in C^f(t)$, then $xI^f y$. Else, $xP^f y$.

This completes the argument that $R^f$ is complete, and hence, a binary relation. Now, we prove that $R^f$ is transitive. Consider $x, y, z \in D$. Consider a type profile $t$, where each column has a value in $\{x, y, z\}$ with at least one column having value $x$, at least one column having value $y$, and at least one column having value $z$ (this is possible since $|A| = m \geq 3$).

We prove transitivity of $P^f$ and $I^f$, and this implies transitivity of $R^f$.

Transitivity of $P^f$: Suppose $xP^f y$ and $yP^f z$. Consider a type profile $t$ such that $t^a = x, t^b = y, t^c = z$. We know that $xP^f y$ implies there exists a type profile $s$ such that $s^a = x$ and $s^b = y$ such that $a \in C^f(s)$ but $b \notin C^f(s)$. By binary independence (Proposition 1), $b \notin C^f(t)$. Similarly, $yP^f z$ implies that there exists a type profile $r$ such that $r^b = y$ and $r^c = z$ such that $b \in C^f(r)$ but $c \notin C^f(r)$. Using binary independence again, $c \notin C^f(t)$. By neutrality, none of the columns in $t$ with values $y$ or $z$ can be in $C^f(t)$. So, one of the columns with value $x$ in $t$ must be in $C^f(t)$. By neutrality again (i.e., if one column with value $x$ is in $C^f(t)$, then every column with value $x$ must be in $C^f(t)$), $a \in C^f(t)$.

This shows that in type profile $t$, where $t^a = x$ and $t^c = z$, we have $a \in C^f(t)$ but $c \notin C^f(t)$. This implies that $xP^f z$.

Transitivity of $I^f$: This proof is similar to the proof of transitivity of $P^f$.

■

## 5 Properties of the Induced Social Welfare Ordering

In this section, we fix an implementable neutral social choice function $f$. We then prove that the social welfare ordering $R^f$ defined in the last section satisfies three specific properties.

Definition 8 *An ordering $R$ on $D$ satisfies* **weak Pareto** *if for all $x, y \in D$ with $x \gg y$ we have $xPy$.*

Definition 9 *An ordering $R$ on $D$ satisfies* **translation invariance (tr-invariance)** *if for all $x, y \in D$ and all $z \in \mathbb{R}^n$ such that $(x+z), (y+z) \in D$ we have $xPy$ implies $(x+z)P(y+z)$ and $xIy$ implies $(x+z)I(y+z)$.*

Definition 10 *An ordering $R$ on $D$ satisfies* **continuity** *if for all $x \in D$, the sets $U^x = \{y \in D : yRx\}$ and $L^x = \{y \in D : xRy\}$ are closed in $D$.*

Proposition 3 **(Axioms for Social Welfare Ordering)** *Suppose $f$ is an implementable and neutral social choice function. Then the social welfare ordering $R^f$ induced by $f$ on $D$ satisfies weak Pareto, tr-invariance, and continuity.*

*Proof*: We show that $R^f$ satisfies each of the properties.

WEAK PARETO: Choose $x, y \in D$ such that $x \gg y$. Start with a profile $t$ where $t^a = y$ for all $a \in A$. Suppose $f(t) = b$. Consider another profile $s = (s^b = x, s^{-b} = t^{-b})$ (i.e. column vector corresponding to $b$ is changed from $y$ to $x$). By PAD, $f(s) = b$ and hence $b \in C^f(s)$. We show that for any $a \neq b$ we have $a \notin C^f(s)$. Choose $\epsilon \gg 0$ but $\epsilon \ll x - y$. By PAD, $f(t^a + \epsilon, s^b = x, t^{-ab}) = b$. Hence, $a \notin C^f(s)$. This shows that $b \in C^f(s)$ but $a \notin C^f(s)$. Hence, by Proposition 2, $xP^f y$.

tr-INVARIANCE: Choose $x, y \in D$ and $z \in \mathbb{R}^n$ such that $(x + z), (y + z) \in D$. We consider two cases.

CASE 1: Suppose $xP^f y$. We show that $(x + z)P^f(y + z)$. Since $xP^f y$, there exists a profile $t = (t^a = x, t^b = y, t^{-ab})$ such that $a \in C^f(t)$ but $b \notin C^f(t)$. Consider the profile $s$, where $s^c = t^c + z$ for all $c \in A$. Fix $\epsilon \gg 0$. Since $a \in C^f(t)$, $f(t^a + \frac{\epsilon}{2}, t^b, t^{-ab}) = a$. Hence, by PAD $f(s^a + \epsilon, s^b, s^{-ab}) = a$. This shows that $a \in C^f(s)$. Since $b \notin C^f(t)$, there is some $\epsilon \gg 0$ such that $f(t^a, t^b + \epsilon, t^{-ab}) \neq b$. We show that $f(s^a, s^b + \frac{\epsilon}{2}, s^{-ab}) \neq b$. Assume for contradiction $f(s^a, s^b + \frac{\epsilon}{2}, s^{-ab}) = b$. By PAD, $f(t^a, t^b + \epsilon, t^{-ab}) = b$. This is a contradiction. Hence, $f(s^a, s^b + \frac{\epsilon}{2}, s^{-ab}) \neq b$. This implies that $b \notin C^f(s)$. Using Proposition 2, $(x + z)P^f(y + z)$.

CASE 2: Suppose $xI^f y$. We show that $(x + z)I^f(y + z)$. Then, there exists a profile $t = (t^a = x, t^b = y, t^{-ab})$ such that $a, b \in C^f(t)$. Consider the profile $s$, where $s^c = t^c + z$ for all $c \in A$. Fix $\epsilon \gg 0$. Since $a \in C^f(t)$, $f(t^a + \frac{\epsilon}{2}, t^b, t^{-ab}) = a$. Hence, by PAD $f(s^a + \epsilon, s^b, s^{-ab}) = a$. This shows that $a \in C^f(s)$. Using an analogous argument, $b \in C^f(s)$. Hence, by Proposition 2, $(x + z)I^f(y + z)$.

CONTINUITY: Fix $x \in D$. We show that the set $U^x = \{y \in D : yR^f x\}$ is closed. Take an infinite sequence $y_1, y_2, \ldots$ such that every point $y_n$ in this sequence satisfies $y_n R^f x$. Let this sequence converge to $z \in D$. Assume for contradiction $xP^f z$. Consider a type profile $t$ such that $t^a = x$ and $t^c = z$ for all $c \neq a$. Since $xP^f z$, we have $c \notin C^f(t)$ for all $c \neq a$. Hence, $C^f(t) = \{a\}$.

Consider $b \neq a$. Since $b \notin C^f(t)$, we know that there exists $\epsilon \gg 0$ and $\epsilon$ arbitrarily close to the zero vector such that $f(t^a, t^b + \epsilon, t^{-ab}) \neq b$. We show that $f(t^a, t^b + \epsilon, t^{-ab}) \neq c$ for all $c \notin \{a, b\}$. Assume for contradiction $f(t^a, t^b + \epsilon, t^c, t^{-abc}) = c$ for some $c \notin \{a, b\}$. Then, by PAD, $f(t^a, t^b, t^c + \epsilon'', t^{-abc}) = c$ for all $\epsilon'' \gg 0$. This implies that $c \in C^f(t)$, which is a contradiction. Hence, $f(t^a, t^b + \epsilon, t^{-ab}) = a$.

This implies that $xR^f(z + \epsilon)$. Since the sequence converges to $z$, there is a point $z' \in D$ arbitrarily close to $z$ such that $z'R^f x$. Since $z$ is arbitrarily close to $z'$, by weak Pareto, $(z + \epsilon)P^f z'$. Using $z'R^f x$, we get $(z + \epsilon)P^f x$. This is a contradiction to the fact that $xR^f(z + \epsilon)$.

To show $L^x = \{y \in D : xR^f y\}$ is closed, take an infinite sequence $y_1, y_2, \ldots$ such that every point $y_n$ in this sequence satisfies $xR^f y_n$. Let this sequence converge to $z$. Assume for contradiction $zP^f x$. Interchanging the role of $x$ and $z$ in the previous argument, we will get $zR^f(x+\epsilon)$ for some $\epsilon \gg 0$. Since the sequence converges to $z$, there is a point $z' \in D$ arbitrarily close to $z$ such that $xR^f z'$. Since $z'$ is arbitrarily close to $z$, $(x+\epsilon)P^f z$ by weak Pareto. This is a contradiction to the fact that $zR^f(x+\epsilon)$. ■

## 6 Multi-dimensional Open Interval Domains

In this section, we prove a representation result. In particular, we prove a proposition related to linear utility representation on open and convex sets.

Proposition **4 (Representation of Social Welfare Ordering)** *Suppose an ordering $R$ on $D$ satisfies weak Pareto, tr-invariance, and continuity. If $D$ is open and convex, then there exists weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ and for all $x, y \in D$*

$$xRy \Leftrightarrow \sum_{i \in N} \lambda_i x_i \geq \sum_{i \in N} \lambda_i y_i.$$

*Proof*: Fix any $z \in D$. Denote $U^z = \{x : xRz\}$, $L^z = \{x : zRx\}$, $D \setminus L^z = \{x : xPz\}$, and $D \setminus U^z = \{x : zPx\}$.

Step 1: We first show that the sets $U^z, L^z, D \setminus U^z$, and $D \setminus L^z$ are convex. We make use of the following fact here.

Fact **1** *Consider a set $X \subseteq D$ and let $X$ satisfy the property that if $x, y \in X$ then $\frac{x+y}{2} \in X$. If $X$ is open in $D$ or closed in $D$, then $X$ is convex.*

The proof of this fact is given in the Appendix A. By continuity, each of the sets $U^z, L^z, D \setminus U^z$, and $D \setminus L^z$ are either open or closed in $D$. Hence, by Fact 1, we only need to verify that these sets are closed under the midpoint operation.

Consider $U^z$. Now, let $x, y \in D$ such that $xRz$ and $yRz$. We will show that $\frac{x+y}{2}Rz$. Note that $\frac{x+y}{2} \in D$ because $D$ is convex. Now, assume for contradiction that $zP\frac{x+y}{2}$. This implies that $xP\frac{x+y}{2}$ and $yP\frac{x+y}{2}$. By tr-invariance, $x + \frac{y-x}{2}P\frac{x+y}{2} + \frac{y-x}{2}$. Hence, $\frac{x+y}{2}Py$. This is a contradiction. Hence, the set $U^z$ is convex.

Similar arguments show that $L^z$, $D \setminus L^z$, and $D \setminus U^z$ are convex.

Step 2: We now show that $z$ is a boundary point of $U^z$. Let $B_\delta(z) = \{x : \|x - z\| < \delta\}$, where $\delta \in \mathbb{R}_{++}$. Since $D$ is open, there exists $\epsilon \gg 0$ such that $(z+\epsilon) \in D \cap B_\delta(z)$ and, by weak Pareto, $(z+\epsilon)Pz$. Further, since $D$ is open, $\epsilon$ can be chosen such that $(z-\epsilon) \in D \cap B_\delta(z)$, and by weak Pareto, $zP(z-\epsilon)$. Hence, for every $\delta > 0$, there exists a point in $B_\delta(z)$ which is in $U^z$ and another point which is not in $U^z$. This shows that $z$ is a boundary point of $U^z$.

Step 3: By the supporting hyperplane theorem, there exists a hyperplane through $z$ supporting the set $U^z$, i.e., there exists a non-zero vector $\lambda \in \mathbb{R}^n \setminus \{0\}$ such that for all $x \in U^z$,

$$\sum_{i=1}^{n} \lambda_i x_i \geq \sum_{i=1}^{n} \lambda_i z_i.$$

Denote the intersection of this hyperplane with the set $D$ as $H^z$.

Step 4: We next show that $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$. Assume for contradiction $\lambda_j < 0$ for some $j \in N$. Since $D$ is open there exists $\epsilon \gg 0$ such that $(z + \epsilon) \in D$. Moreover, we can choose $\epsilon$ such that

$$\sum_{i=1}^{n} \lambda_i \epsilon_i < 0.$$

By weak Pareto $(z + \epsilon)Pz$. Hence, $(z + \epsilon) \in U^z$. Thus,

$$\sum_{i=1}^{n} \lambda_i (z_i + \epsilon_i) \geq \sum_{i=1}^{n} \lambda_i z_i.$$

This implies that

$$\sum_{i=1}^{n} \lambda_i \epsilon_i \geq 0.$$

This is a contradiction. Hence, $\lambda_i \geq 0$ for all $i \in N$.

Step 5: Now, consider $x \in D$ such that

$$\sum_{i=1}^{n} \lambda_i x_i > \sum_{i=1}^{n} \lambda_i z_i.$$

We will show that $xPz$. Assume for contradiction $zRx$. We consider two cases.

Case 1: Suppose $zPx$. Since $D$ is open, there exists a point $z'$ in $B_\delta(z)$ for some $\delta \in \mathbb{R}_+$ such that

a) $z$ lies on the line segment joining $z'$ and $x$ and

b) $x$ and $z'$ lie on opposite sides of the hyperplane $H_z$, i.e.,

$$\sum_{i=1}^{n} \lambda_i z'_i < \sum_{i=1}^{n} \lambda_i z_i.$$

By (b) and using Step 3, $zPz'$. By our assumption $zPx$. Hence, $x, z' \in D \setminus U^z$. By Step 1, $D \setminus U^z$ is convex. Since $z$ is in the convex hull of $x$ and $z'$, we get that $zPz$. This is a contradiction.

CASE 2: Suppose $zIx$. Since $D$ is open, there exists $x' = x - \epsilon$ for some $\epsilon \gg 0$ such that

$$\sum_{i=1}^{n} \lambda_i x'_i > \sum_{i=1}^{n} \lambda_i z_i.$$

By weak Pareto $xPx'$. Hence, $zPx'$. By Case 1, this is not possible. This is a contradiction.

Hence, in both cases we reach a contradiction, and conclude that $xPz$.

STEP 6: Now, consider $x \in D$ such that

$$\sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{n} \lambda_i z_i.$$

We will show that $xIz$. Suppose not. There are two cases to consider.

CASE 1: Assume for contradiction $xPz$. By continuity, the set $\{y : yPz\}$ is open in $D$. Since $D$ is open in $\mathbb{R}^n$, we get that $\{y : yPz\}$ is open in $\mathbb{R}^n$. Hence, there exists $\delta \in \mathbb{R}_+$ such that for every point $x' \in B_\delta(x)$ we have $x'Pz$. Choose $\epsilon \gg 0$ such that for $x'' = x - \epsilon$ we have $x'' \in B_\delta(x)$. Hence, $x''Pz$. By Step 4, $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$. Hence, we get

$$\sum_{i=1}^{n} \lambda_i x''_i < \sum_{i=1}^{n} \lambda_i z_i.$$

But this is a contradiction since $x''Pz$ implies $x'' \in U^z$, which in turn implies that

$$\sum_{i=1}^{n} \lambda_i x''_i \geq \sum_{i=1}^{n} \lambda_i z_i.$$

CASE 2: Assume for contradiction $zPx$. By continuity, the set $\{y : zPy\}$ is open in $D$. Hence, there exists $\delta \in \mathbb{R}_+$ such that for every point in $x' \in B_\delta(x)$ we have $zPx'$. Choose $\epsilon \gg 0$ such that for $x'' = x + \epsilon$ we have $x'' \in B_\delta(x)$. Hence, $zPx''$. By Step 4, $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$. Hence, we get

$$\sum_{i=1}^{n} \lambda_i x''_i > \sum_{i=1}^{n} \lambda_i z_i.$$

By Step 5, this implies that $x''Pz$. This is a contradiction.

This shows that for any $z$, there exists $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ such that for all $x \in D$, we have

$$xRz \Leftrightarrow \sum_{i=1}^{n} \lambda_i x_i \geq \sum_{i=1}^{n} \lambda_i z_i.$$

In other words, $H^z$ contains all the points in $D$ which are indifferent to $z$ under $R$. Moreover, on one side of $H^z$ we have points in $D$ which are better than $z$ under $R$ and on the other side, we have points which are worse than $z$ under $R$.

Finally, pick any two points $x$ and $y$ in $D$. Since $D$ is open and convex, we can connect $x$ and $y$ by a series of intersecting open balls along the convex hull of $x$ and $y$, with each of these open balls contained in $D$. By tr-invariance, for any two points $x'$ and $y'$ in such an open ball, $H^{x'}$ and $H^{y'}$ have to be parallel to each other [6]. Since such open balls intersect each other, the hyperplanes $H^x$ and $H^y$ are parallel to each other. This completes the proof. ■

When $D = \mathbb{R}^n$, this result is well known due to Blackwell and Grishick (1954) (see also recent proofs in the utility representation literature - d'Aspremont and Gevers (1977), Blackorby et al. (1984), Trockel (1992), and d'Aspremont and Gevers (2002)).

We are now ready to state the proof of Theorem 3.

PROOF OF THEOREM 3

*Proof*: Suppose $f$ is neutral and implementable. Note that since for every $i \in N$, $T_i$ is an open interval domain, then $D$ must be convex and open in $\mathbb{R}^n$ - indeed, $D$ is an open rectangle in $\mathbb{R}^n$. Hence, by Proposition 2, a neutral and implementable SCF $f$ induces a social welfare ordering $R^f$ on $D$. By Proposition 3, $R^f$ satisfies continuity, weak Pareto, and tr-invariance. By Proposition 4 (since $D$ is open and convex), there exists weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ such that for every $x, y \in D$ we have

$$xR^f y \Leftrightarrow \sum_{i \in N} \lambda_i x_i \geq \sum_{i \in N} \lambda_i y_i.$$

Finally, by Lemma 2 for all $t \in D^m$, $f(t) \in C^f(t)$. Hence, $t^{f(t)} R^f t^b$ for all $b \in A$ and for all $t \in D^m$. ■

## 6.1 Discussions

In this section, we make several observations relating to our main result.

[6] To see this, assume for contradiction $H^{x'}$ and $H^{y'}$ are not parallel. Then, one can find a point $x'' \neq x'$ on $H^{x'}$ such that $x'' + (y' - x') \notin H^{y'}$ and $[x'' + (y' - x')] \in D$. This implies either $[x'' + (y' - x')Py']$ or $y'P[x'' + (y' - x')]$. By tr-invariance, $x''Ix'$ implies $[x'' + (y' - x')]Iy'$, which is a contradiction.

Affine Maximizer and Weighted Welfare Maximizer Domains. A plausible conjecture is that every domain where neutral and implementable social choice functions are weighted welfare maximizers are also domains where implementable social choice functions are affine maximizers. This conjecture is false. To see this, observe the domain in Example 1. The domain in this example, $(0,1)^2$, is a 2-dimensional open interval domain. However we have already seen that it admits implementable social choice functions that are non-affine-maximizers (of course, these social choice functions are not neutral). This observation emphasizes the fact that neutrality plays a critical role in our result.

Auction domains are not covered. It is well known that in auction domains, there are social choice functions other than affine maximizers that are implementable. Concrete examples can be found in Lavi et al. (2003) and Dobzinski and Nisan (2011). These social choice functions are also neutral. Hence, in auction domains, there are neutral social choice functions which are implementable, but not weighted welfare maximizers. It can be reconciled with our result in the following way. Auction domains are restricted domains which are not necessarily open (or even full dimensional). For example, consider the sale of two objects to two buyers. The set of allocations can be $\{a, b, c, d\}$, where $a$ denotes buyer 1 gets both the objects, $b$ denotes buyer 2 gets both the objects, $c$ denotes buyer 1 gets object 1 and buyer 2 gets object 2, and $d$ denotes buyer 1 gets object 2 and buyer 2 gets object 1. Note here that in every utility vector $t^a$ for allocation $a$ buyer 2 will have zero valuation. Similarly, in every utility vector $t^b$ for allocation $b$ buyer 1 will have zero valuation. Hence, this domain is not open.

No ordering without neutrality. If we drop neutrality and replace it with non-imposition, then Roberts' Theorem says that affine maximizers (as in Theorem 1) are the only implementable social choice functions. But affine maximizers do not necessarily induce the ordering we discussed. This is because of the $\kappa(\cdot)$ terms in the affine maximizers. For example, consider a type profile $t = (t^a = x, t^b = y, t^{-ab})$. Suppose $a \in C^f(t)$ but $b \notin C^f(t)$. Here, the $\kappa(a)$ term may be higher than $\kappa(b)$ such that when we permute the columns of $a$ and $b$ and get the new type profile $s = (s^a = y, s^b = x, t^{-ab})$, we still have $a \in C^f(s)$ and $b \notin C^f(s)$. Thus, our social welfare ordering is not induced here.

Anonymity gives efficiency. Consider the following additional condition on every social choice function.

Definition **11** *A social choice function $f$ is* **anonymous** *if for every $t \in T$ and every permutation $\sigma$ on the row vectors (agents) of $t$, we have $f(\sigma(t)) = f(t)$.*

Definition **12** *An ordering $R$ on $D$ satisfies* **anonymity** *if for every $x, y \in D$ and every permutation $\sigma$ on agents we have $xIy$ if $x = \sigma(y)$.*

Lemma **5** *Suppose $f$ is implementable and anonymous. Then, $R^f$ satisfies anonymity.*

*Proof*: Let $\sigma$ be a permutation of the set of agents. For any vector $x \in D$, we write $\bar{\sigma}(x)$ to denote the permutation of vector $x$ induced by the permutation $\sigma$ on sets of agents. Consider $x, y \in D$ such that $y = \bar{\sigma}(x)$. Assume for contradiction $xP^f y$. Consider a type profile $t$ such that $t^a = x$ and $t^b = y$ for all $b \neq a$. Hence, $C^f(t) = \{a\} = f(t)$. Let $s$ be the type profile such that $s^c = \bar{\sigma}(t^c)$ for all $c \in A$. Since $f$ is anonymous $f(s) = a$. Hence, $yR^f \bar{\sigma}(y)$, which futher implies that $xP^f \bar{\sigma}(\bar{\sigma}(x))$. Repeating this argument again, we will get $\bar{\sigma}(y)R^f \bar{\sigma}(\bar{\sigma}(y))$. Hence, $xP^f \bar{\sigma}(\bar{\sigma}(\bar{\sigma}(x)))$. Clearly, after repeating this procedure some finite number of times, we will be able to conclude $xP^f x$, which is a contradiction. ■

It is straightforward to show using Theorem 3 that every implementable, neutral, and anonymous social choice function in an open interval domain is the efficient social choice function. Here, we show that this result holds for some other domains too. The proof is an adaptation of an elegant proof by Milnor (1954) (see also Theorem 4.4 in d'Aspremont and Gevers (2002)). We give the proof in Appendix A.

Theorem **4** *Suppose $f$ is implementable, neutral, and anonymous. If $T = [0, H)^{m \times n}$, where $H \in \mathbb{R} \cup \{\infty\}$, then $f$ is the efficient social choice function.*

Note here that the domain in Theorem 4 always includes the origin (this is crucial for the proof) and is not open from "left". Hence, this result is not a corollary to Theorem 3.

## 7 Roberts' Affine-Maximizer Theorem

In this section, we prove the general version of Roberts' Theorem using the version of Roberts' Theorem with neutrality, which we have proved earlier. We assume throughout that the domain is **unrestricted**, i.e., $T = \mathbb{R}^{m \times n}$. Although our proof of the general Roberts' Theorem uses elements developed in earlier proofs, we believe nonetheless that it offers some new insights into the result. The main idea behind our proof is to transform an arbitrary implementable social choice function to a neutral implementable social choice function. Then, we can readily use Roberts' theorem with neutrality on the new social choice function to get the Roberts' Theorem.

Consider a mapping $\delta : A \rightarrow \mathbb{R}$. Denote $1_{\delta(a)}$ as the vector of $\delta(a)$ s in $\mathbb{R}^n$. Let $1_\delta \equiv (1_{\delta(a)}, 1_{\delta(b)}, \ldots)$ be the profile of $m$ such vectors, each corresponding to an allocation in $A$. For any social choice function $f$, define $f^\delta$ as follows. For every $t \in T$, let $(t + 1_\delta) \in T$ be such that $(t + 1_\delta)^a = t^a + 1_{\delta(a)}$ for all $a \in A$. For every $t \in T$, let

$$f^\delta(t) = f(t + 1_\delta).$$

Since $\delta(a)$ is finite for all $a \in A$, the social choice function $f^\delta$ is well-defined.

Proposition **5 (Implementability Invariance)** *For every $\delta : A \rightarrow \mathbb{R}$, if $f$ is implementable, then $f^\delta$ is implementable.*

*Proof*: Since $f$ is implementable, there exists a payment function $p$ which implements it. We define another payment function $p^\delta$ as follows. For every $t \in T$ and every $i \in N$,

$$p_i^\delta(t) = p_i(t + 1_\delta) + \delta(f^\delta(t)).$$

We will show that $p^\delta$ implements $f^\delta$. To see this, fix an agent $i \in N$ and $t_{-i} \in \mathbb{T}_{-i}$. Let $s = (s_i, t_{-i})$ and note the following.

$$\begin{aligned} t_i^{f^\delta(t)} + p_i^\delta(t) &= t_i^{f(t+1_\delta)} + p_i(t + 1_\delta) + \delta(f^\delta(t)) \\ &= t_i^{f(t+1_\delta)} + p_i(t + 1_\delta) + \delta(f(t + 1_\delta)) \\ &= (t + 1_\delta)_i^{f(t+1_\delta)} + p_i(t + 1_\delta) \\ &\geq (t + 1_\delta)_i^{f(s+1_\delta)} + p_i(s + 1_\delta) \\ &= (t + 1_\delta)_i^{f^\delta(s)} + p_i(s + 1_\delta) \\ &= t_i^{f^\delta(s)} + \delta(f^\delta(s)) + p_i(s + 1_\delta) \\ &= t_i^{f^\delta(s)} + p^\delta(s), \end{aligned}$$

where the inequality followed from the implementability of $f$ by $p$. Hence, $p^\delta$ implements $f^\delta$. ■

Our next step is to find a mapping $\delta : A \to \mathbb{R}$ such that $f^\delta$ is neutral. We will need the following property of choice sets.

LEMMA **6** *Suppose $f$ is implementable and satisfies non-imposition. Let $t$ be a type profile such that $C^f(t) = \{a\}$ for some $a \in A$. Then, for some $\epsilon \in \mathbb{R}^n_{++}$, $a \in C^f(s)$, where $s^a = t^a - \epsilon$ and $s^b = t^b$ for all $b \neq a$.*

*Proof*: Since $C^f(t) = \{a\}$, we have $f(t) = a$ (by Lemma 2). Choose some $b \neq a$. Since $b \notin C^f(t)$, there exists $\epsilon_b \in \mathbb{R}^n_{++}$ such that $b \notin C^f(u)$, where $u^b = t^b + \epsilon_b$ and $u^c = t^c$ for all $c \neq b$. Indeed, by Proposition 1, $C^f(u) = \{a\}$. Now, consider the type profile $v$ such that $v^c = t^c + \epsilon_c$ for all $c \neq a$ and $v^a = t^a$. We will show that $C^f(v) = \{a\}$.

To show this, we go from $t$ to $v$ in $(m-1)$ steps. In the first step, we choose an arbitrary allocation $b \neq a$, and consider a type profile $x$, where $x^b = v^b$ and $x^c = t^c$ for all $c \neq b$. By definition of $\epsilon_b$, we have $C^f(x) = \{a\}$. Next, we choose another allocation $c \notin \{a, b\}$, and consider a type profile $y$ such that $y^d = x^d$ if $d \neq c$ and $y^d = v^d$ otherwise. We first show that $c \notin C^f(y)$. Assume for contradiction, $c \in C^f(y)$, then by PAD, $c \in C^f(x)$. This is a contradiction. Hence, $c \notin C^f(y)$, and by Proposition 1, $C^f(y) = \{a\}$. We now repeat this procedure by choosing $d \notin \{a, b, c\}$ and considering a type profile $z$ where utility vector of $z$ is increased to $v^d$ and every other utility vector remains at $y$. After a finite steps, we will reach the type profile $v$ with $C^f(v) = \{a\}$.

Now, choose $\epsilon = \frac{1}{2}\min_{b \neq a} \epsilon_b$. Consider a type profile $s$ such that $s^b = t^b$ for all $b \neq a$ and $s^a = t^a - \epsilon$. By PAD (from $v$ to $s$), $a \in C^f(s)$. ■

Now, we define a set which can also be found in Roberts' original proof (see also Lavi et al. (2009)). For every $a, b \in A$ and every social choice function $f$ define the $P$-set as

$$P^f(a,b) = \{\alpha \in \mathbb{R}^n : \exists t \in T \text{ such that } a \in C^f(t), t^a - t^b = \alpha\}.$$

Roberts (1979) and Lavi et al. (2009) define the $P$-set slightly differently. They let $P^f(a,b) = \{\alpha \in \mathbb{R}^n : \exists t \in T \text{ such that } f(t) = a, t^a - t^b = \alpha\}$. Our notion of $P$-set is the interior of the $P$-set they define.

The $P$-sets are non-empty if the social choice function satisfies non-imposition. To see this, choose $a, b \in A$ and a social choice function $f$. By non-imposition, there must exist a $t \in T$ such that $f(t) = a$, which implies that $a \in C^f(t)$ and $(t^a - t^b) \in P^f(a,b)$.

We want to characterize a neutral social choice function by the properties of its $P$-sets. Here is a necessary and sufficient condition.

Proposition **6 (Neutrality)** *Suppose $f$ is an implementable social choice function. The social choice function $f$ is neutral if and only if $P^f(a,b) = P^f(c,d)$ for all $a,b,c,d \in A$.*

*Proof*: Suppose $f$ is implementable and neutral. Let $\alpha \in P^f(a,b)$. So, for some type profile $t$, we have $a \in C^f(t)$ and $t^a - t^b = \alpha$. Now, permuting $a, b$ respectively with $c, d$, we get a new type profile $s$ with $s^c = t^a, s^d = t^b, s^a = t^c, s^b = t^d$. By neutrality, $c \in C^f(s)$ and $s^c - s^d = t^a - t^b = \alpha$. So, $\alpha \in P^f(c,d)$. Exchanging the role of $(a,b)$ and $(c,d)$ in this argument, we get that $\alpha \in P^f(c,d)$ implies $\alpha \in P^f(a,b)$. Thus, $P^f(a,b) = P^f(c,d)$.

Now, suppose that $f$ is implementable and $P^f(a,b) = P^f(c,d)$ for all $a,b,c,d \in A$. Consider a permutation $\rho$ of $A$. Without loss of generality, assume that $\rho$ is a transposition, i.e., for some $a, b \in A$ we have $\rho(a) = b, \rho(b) = a$, and $\rho(c) = c$ for all $c \notin \{a,b\}$. Consider a type profile $t \in T$ and let $s$ be the type profile induced by permutation $\rho$ on $t$, i.e., $s^a = t^b, s^b = t^a$, and $s^{-ab} = t^{-ab}$. We show $f$ is neutral in several steps.

Step 1: Suppose $a \notin C^f(t)$. We show that $b \notin C^f(s)$. Assume for contradiction $b \in C^f(s)$. Let $c \in C^f(t)$. Such a $c$ exists since $C^f(t)$ is non-empty. Note that $c \neq a$. There are two cases to consider.

Case 1: Suppose $c = b$. Because, $b \in C^f(s)$, we get that $(t^a - t^c) \in P^f(b,a) = P^f(a,c)$.

Case 2: Suppose $c \notin \{a,b\}$. Again, because $b \in C^f(s)$, we get that $(t^a - t^c) \in P^f(b,c) = P^f(a,c)$.

So, we get $(t^a - t^c) \in P^f(a,c)$ in both the cases. Then for some $\epsilon \in \mathbb{R}^n$ and some type profile $v = (v^a = t^a + \epsilon, v^c = t^c + \epsilon, v^{-ac})$, we have $a \in C^f(v)$. Consider the type profile $u$ such that $u^a = t^a, u^c = t^c$, and $u^d = v^d - \epsilon$ for all $d \notin \{a,c\}$. By PAD, $a \in C^f(u)$. But, in both $t$ and $u$, the utility vectors corresponding to $a$ and $c$ are respectively $t^a$ and $t^c$. Since

$a \notin C^f(t)$ and $c \in C^f(t)$, by Proposition 1, $a \notin C^f(u)$. This is a contradiction.

STEP 2: Suppose $a \in C^f(t)$. We show that $b \in C^f(s)$. Assume for contradiction $b \notin C^f(s)$. By Step 1, $a \notin C^f(t)$. This is a contradiction.

STEP 3: Suppose $c \in C^f(t)$, where $c \notin \{a, b\}$. We show that $c \in C^f(s)$. Since $c \in C^f(t)$, we have $(t^c - t^a), (t^c - t^b) \in P^f(c, b)$. Assume for contradiction $c \notin C^f(s)$. Then, for some $d \neq c$, we have $d \in C^f(s)$. There are two cases to consider.

CASE 1: Suppose $d \notin \{a, b, c\}$. In that case, by Proposition 1 (applied to $s$ and $t$), $c \notin C^f(t)$. This is a contradiction.

CASE 2: Suppose $d \in \{a, b\}$. Without loss of generality, let $d = a$. So, $a \in C^f(s)$ but $c \notin C^f(s)$. Now, since $(t^c - t^b) \in P^f(c, b) = P^f(c, a)$, there exists a type profile $u = (u^a = t^b + \epsilon, u^c = t^c + \epsilon, u^{-ac})$ such that $c \in C^f(u)$. By PAD, $c \in C^f(v)$, where $v^a = t^b, v^c = t^c$, and $v^d = u^d - \epsilon$ for all $d \notin \{a, c\}$. By Proposition 1, we get that if $c \notin C^f(s)$, then $a \notin C^f(s)$. This is a contradiction.

STEP 4: Suppose $c \notin C^f(t)$. Assume for contradiction $c \in C^f(s)$. Exchanging the role of $s$ and $t$ in Step 3, we get that $c \in C^f(t)$. This is a contradiction.

Combining all the steps, we get that $C^f(s) = \{\rho(c) : c \in C^f(t)\}$, i.e., $f$ is neutral. ■

We begin by noting two properties of the $P$-sets. Identical properties have been established in (Lavi et al., 2009; Vohra, 2008) for their version of $P$-sets. We give proofs which are also more direct.

LEMMA 7 *Suppose $f$ is implementable and satisfies non-imposition. The following statements are true for every $a, b, c \in A$.*

1. *If $(\beta - \epsilon) \in P^f(a, b)$ for some $\beta \in \mathbb{R}^n$ and some $\epsilon \in \mathbb{R}^n_{++}$, then $-\beta \notin P^f(b, a)$.*
2. *If $\beta \in P^f(a, b)$ and $\alpha \in P^f(b, c)$, then $(\beta + \alpha) \in P^f(a, c)$.*

*Proof*: Fix $a, b, c \in A$.

PROOF OF (1): Suppose $(\beta - \epsilon) \in P^f(a, b)$ for some $\beta \in \mathbb{R}^n$ and some $\epsilon \in \mathbb{R}^n_{++}$. Assume for contradiction that $-\beta \in P^f(b, a)$. So, there exists some type profile $t$ such that $b \in C^f(t)$ and $t^a - t^b = \beta$. Consider the type profile $s$ such that $s^a = t^a - \epsilon$ and $s^c = t^c$ for all $c \neq a$. Note that $(s^a - s^b) = (\beta - \epsilon)$. We first show that $a \in C^f(s)$. Since $(\beta - \epsilon) \in P^f(a, b)$, there is some profile $u = (u^a = s^a + \alpha, u^b = s^b + \alpha, u^{-ab})$, where $\alpha \in \mathbb{R}^n$, such that $a \in C^f(u)$. By PAD, there is a profile $v = (v^a = s^a, v^b = s^b, v^{-ab})$ such that $a \in C^f(v)$. We consider two cases.

Case 1: Suppose $b \notin C^f(v)$. Then, by Proposition 1, $b \notin C^f(s)$. By PAD, $b \notin C^f(t)$, which is a contradiction.

Case 1: Suppose $b \in C^f(v)$. Then, by Proposition 1, $a \in C^f(s)$ if and only if $b \in C^f(s)$. If $b \notin C^f(s)$, as in Case 1, we have a contradiction due to PAD. Hence, $a, b \in C^f(s)$. Consider the type profile $x$ such that $x^a = t^a$, $x^b = t^b + \frac{\epsilon}{2}$, and $x^c = t^c$ for all $c \notin \{a, b\}$. By PAD, $f(x) = a$. Hence, $b \notin C^f(t)$. This is a contradiction.

Proof of (2): Suppose $\beta \in P^f(a, b)$ and $\alpha \in P^f(b, c)$. Then, there must exist $t \in T$ such that $a \in C^f(t)$ and $t^a - t^b = \beta$. Now, consider a type profile $s$ such that $s^a = t^a$, $s^b = t^b$, $s^c = t^b - \alpha$, and $s^d$ is sufficiently low for all $d \notin \{a, b, c\}$. We show that for all $d \notin \{a, b, c\}$, we have $d \notin C^f(s)$. Assume for contradiction $d \in C^f(s)$. Then, by PAD, $a \notin C^f(t)$, which is a contradiction. So, $C^f(s) \subseteq \{a, b, c\}$.

We show that $a \in C^f(s)$. Assume for contradiction $a \notin C^f(s)$. Then, by Proposition 1, $b \notin C^f(s)$. This implies that $C^f(s) = \{c\}$ (by Lemma 2). By Lemma 6, $(-\alpha - \epsilon) \in P^f(c, b)$. By (1), $\alpha \notin P^f(b, c)$. This is a contradiction.

This implies that $a \in C^f(s)$, and hence, $(s^a - s^c) \in P^f(a, c)$. But $s^a - s^c = t^a - t^b + \alpha = \beta + \alpha$ implies that $(\beta + \alpha) \in P^f(a, c)$. ■

We are now ready to define the mapping that will make any social choice function neutral. Define the following mapping $\kappa : A \to \mathbb{R}$ as follows. For all $a \in C^f(0)$ [7], let $\kappa(a) = 0$. For all $a \notin C^f(0)$, define $\kappa(a)$ as follows. Denote a type vector $t$ as $1^b_\epsilon$, where all utility (column) vectors except one, say $t^b$, is zero vector and $t^b = 1_\epsilon$ for some $\epsilon \in \mathbb{R}$. For all $a \notin C^f(0)$,

$$\kappa(a) = \{\epsilon \in \mathbb{R}_+ : C^f(1^a_\epsilon) = C^f(0) \cup \{a\}\}.$$

Our first claim is that for all $a \in A$, $\kappa(a) \in \mathbb{R}_+$ exists.

Lemma 8 *Suppose $f$ is implementable and satisfies non-imposition. Then, for all $a \in A$, $\kappa(a) \in \mathbb{R}_+$ and is unique. Moreover, $\kappa(a) = \inf\{\epsilon \in \mathbb{R}_+ : a \in C^f(1^a_\epsilon)\}$.*

*Proof*: For all $a \in C^f(0)$, $\kappa(a) = 0$, and hence, the lemma is true. Consider $a \notin C^f(0)$. If $\kappa(a)$ exists, by PAD, it is unique. We show that $\kappa(a)$ exists. We do this in two steps.

Step 1: We show that there exists an $\epsilon \in \mathbb{R}_+$ such that $a \in C^f(1^a_\epsilon)$. By non-imposition, there exists a type profile $t$ such that $f(t) = a$. By PAD, there exists an $\epsilon \in \mathbb{R}$ such that $a \in C^f(1^a_\epsilon)$. Moreover $\epsilon > 0$ since $a \notin C^f(0)$.

[7] Here, 0 denotes the type profile, where every agent's type is the $m$-dimensional zero vector.

Step 2: We now prove the lemma. Define

$$\kappa(a) = \inf\{\epsilon : a \in C^f(1^a_\epsilon)\}.$$

By Step 1, $\kappa(a)$ exists. We show that $C^f(1^a_{\kappa(a)}) = C^f(0) \cup \{a\}$. Consider $b \notin (C^f(0) \cup \{a\})$. By PAD, if $b \in C^f(1^a_{\kappa(a)})$, then $b \in C^f(0)$, which is a contradiction. Hence, $b \notin C^f(1^a_{\kappa(a)})$. Next, by Proposition 1, we can conclude that either $C^f(1^a_{\kappa(a)}) = C^f(0) \cup \{a\}$ or $C^f(1^a_{\kappa(a)}) = \{a\}$. Assume for contradiction $C^f(1^a_{\kappa(a)}) = \{a\}$. Then, by Lemma 6, there exists $\epsilon \in \mathbb{R}^n_{++}$ such that $a \in C^f(1^a_{\kappa(a)-\epsilon})$. This is a contradiction by the definition of $\kappa(a)$. This shows that $C^f(1^a_{\kappa(a)}) = C^f(0) \cup \{a\}$. ■

We now prove a critical lemma.

Lemma 9 *Suppose $f$ is implementable and satisfies non-imposition. Let $t$ be a type profile such that $t^a = 1_{\kappa(a)}$ for all $a \in A$. Then, $C^f(t) = A$.*

*Proof*: We start from the type profile 0 and move to $t$ in finite number of steps. Consider a set $A^0 \subseteq A$. Initially, $A^0 = A \setminus C^f(0)$. Now, choose $a \in A^0$, and consider $1^a_{\kappa(a)}$. By definition of $\kappa(a)$, $C^f(1^a_{\kappa(a)}) = \{a\} \cup C^f(0)$. Now, set $A^0 := A^0 \setminus \{a\}$, and choose $b \in A^0$. We now define a type profile $s$ such that $s^a = 1_{\kappa(a)}$ and $s^b = 1_{\kappa(b)}$ but $s^c = 0$ for all $c \notin \{a, b\}$. By Proposition 1, either $C^f(s) = C^f(1^a_{\kappa(a)}) \cup \{b\}$ or $C^f(s) = \{b\}$. The latter case is not possible by Lemma 6 since it will imply $b \in C^f(1^b_{\kappa(b)-\epsilon})$ for some $\epsilon \in \mathbb{R}^n_{++}$, which will violate the definition of $\kappa(b)$. Hence, $C^f(s) = C^f(1^a_{\kappa(a)}) \cup \{b\}$. Now, we set $A^0 := A^0 \setminus \{b\}$, and repeat. Since $A$ is finite, this process will terminate with type profile $t$ such that $C^f(t) = A$. ■

We now have all ingredients for proving the Roberts' Theorem.

Theorem 5 (Roberts (1979)) *Suppose $T = \mathbb{R}^{m \times n}$. If $f$ is an implementable social choice function and satisfies non-imposition, then there exist weights $\lambda \in \mathbb{R}^n_+ \setminus \{0\}$ and a function $\kappa : A \to \mathbb{R}$ such that for all $t \in T$,*

$$f(t) \in \arg\max_{a \in A} \Big[\sum_{i \in N} \lambda_i t^a_i - \kappa(a)\Big]$$

*Proof*: Since $f$ is implementable and satisfies non-imposition, by Lemma 8, there exists a mapping $\kappa : A \to \mathbb{R}$ satisfying properties stated in Lemmas 8 and 9. Now, consider the social choice function $f^\kappa$. By Proposition 5, $f^\kappa$ is implementable. By definition, $f^\kappa(0) = f(1_\kappa)$. By Lemma 9, $C^{f^\kappa}(0) = C^f(t) = A$. This implies that $0 \in P^{f^\kappa}(a, b)$ for all $a, b \in A$.

Now, pick $a, b, c, d \in A$ and let $\beta \in P^{f^\kappa}(a, b)$. But $0 \in P^{f^\kappa}(b, d)$. By Lemma 7, $\beta \in P^{f^\kappa}(a, d)$. Now, using $0 \in P^{f^\kappa}(c, a)$, and applying Lemma 7 again, we get $\beta \in P^{f^\kappa}(c, d)$. By Proposition 6, $f^\kappa$ is neutral. By Theorem 3, $f^\kappa$ is a weighted welfare maximizer. This implies that there exists $\lambda \in \mathbb{R}^n \setminus \{0\}$ such that for every $t \in T$,

$$f^\kappa(t) \in \arg\max_{a \in A} \sum_{i=1}^{n} \lambda_i t^a_i.$$

But this implies that, for every $t \in T$,

$$f^{\kappa}(t - 1_{\kappa}) \in \arg\max_{a \in A} \sum_{i=1}^{n} \lambda_i (t - 1_{\kappa})_i^a.$$

This in turn implies that, for every $t \in T$,

$$f(t) \in \arg\max_{a \in A} \big[ \sum_{i=1}^{n} \lambda_i [t_i^a - \kappa(a)] \big]$$

Since we can assume without loss of generality that $\lambda_i \in [0, 1]$ for all $i \in N$, we can immediately infer Roberts' Theorem. ■

To summarize, Roberts' Theorem can be proved using Roberts' Theorem with neutrality by transforming any social choice function to a neutral social choice function as given by Proposition 6. This transformation seems to require that the domain be unrestricted.

## 8 Conclusion

We have shown that every implementable and neutral social choice function is a weighted welfare maximizer in a bounded domain. Our proof technique reduces the problem of characterizing such social choice functions to the problem of characterizing orderings over Euclidean space, a problem which has been studied at length in social choice theory. An open question is whether this approach is useful more generally, i.e., whether it can be deployed in auction domains.

Finally, we show how Roberts' Theorem (the general version) can be proved using Roberts' Theorem with neutrality. This proof requires transforming any implementable social choice function into a neutral and implementable social choice function. To our knowledge, this transformation seems to require the unrestricted domain.

We summarize our main contribution in Figure 1. The arrows in this figure indicate implications. As the figure shows, our results can be thought to be equivalence of the PAD condition and implementability in the presence of neutrality in open interval domains. It will be interesting to investigate this equivalence in the absence of neutrality.

## References


Archer, A. and R. Kleinberg (2008): "Truthful Germs are Contagious: A Local to Global Characterization of Truthfulness," in *In Proceedings of the 9th ACM conference on Electronic commerce (EC-08)*, Springer (Lecture Notes in Computer Science).

Ashlagi, I., M. Braverman, A. Hassidim, and D. Monderer (2010): "Monotonicity and Implementability," *Econometrica*, 78, 1749–1772.

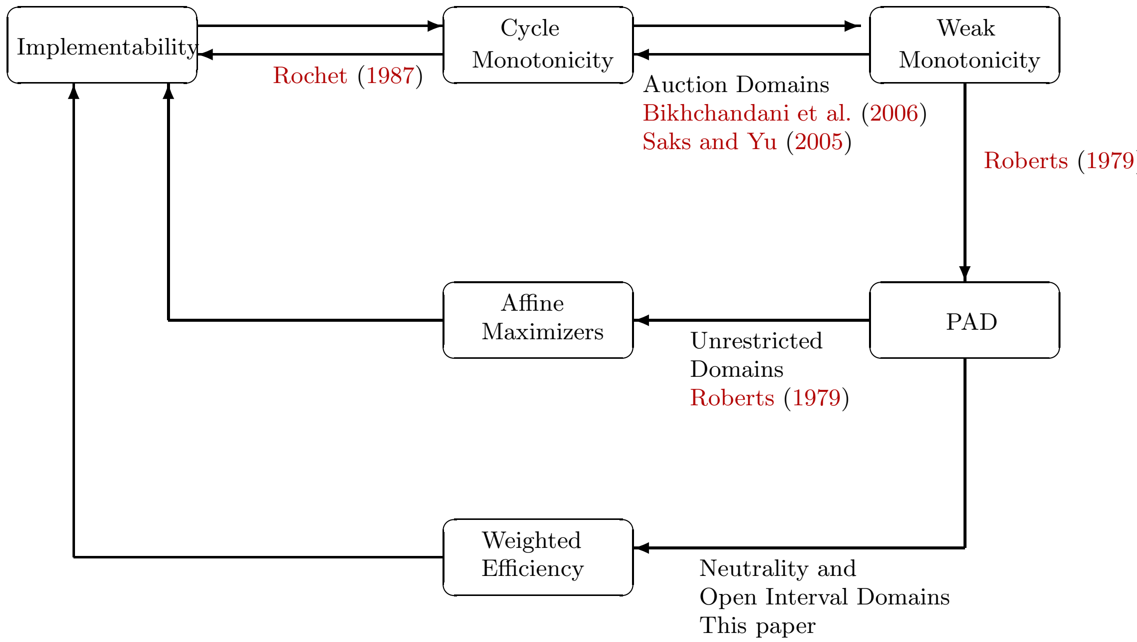


Figure 1: Understanding Implementability

Berger, A., R. Muller, and S. H. Naeemi (2010): "Path-Monotonicity and Incentive Compatibility," Working Paper, Maastricht University.

Bikhchandani, S., S. Chatterji, R. Lavi, A. Mualem, N. Nisan, and A. Sen (2006): "Weak Monotonicity Characterizes Deterministic Dominant Strategy Implementation," *Econometrica*, 74, 1109–1132.

Blackorby, C., D. Donaldson, and J. A. Weymark (1984): "Social Choice with Interpersonal Utility Comparisons: A Diagrammatic Introduction," *International Economic Review*, 25, 327–356.

Blackwell, D. and M. A. Grishick (1954): *Theory of Games and Statistical Decisions*, Wiley, New York.

Carbajal, J. C., A. McLennan, and R. Tourky (2011): "Truthful Implementation and Aggregation in Restricted Domains," Working Paper, University of Queensland.

Chung, K.-S. and J. C. Ely (2006): "Ex-Post Incentive Compatible Mechanism Design," Working Paper, University of Minnesota, http://sites.google.com/site/kimsauchung/researchpapers.

Clarke, E. (1971): "Multipart Pricing of Public Goods," *Public Choice*, 11, 17–33.

d'Aspremont, C. and L. Gevers (1977): "Equity and the Informational Basis of Collective Choice," *Review of Economic Studies*, 44, 199–209.

——— (2002): *Handbook of Social Choice and Welfare - Volume 2*, Elsevier, Amsterdam, chap. Social Welfare Functionals and Interpersonal Comparability (Chapter 10), 459–541, ediors: Kenneth J. Arrow, Amartya K. Sen, and K. Suzumura.

Dobzinski, S. and N. Nisan (2009): "A Modular Approach to Roberts' Theorem," in *In Proceedings of the 2nd International Symposium on Algorithmic Game Theory (SAGT 2009)*, Springer (Lecture Notes in Computer Science).

——— (2011): "Multi-Unit Auctions: Beyond Roberts," in *Forthcoming, Proceedings of* $12^{th}$ *ACM Conference on Electronic Commerce*, ACM Press.

Groves, T. (1973): "Incentives in Teams," *Econometrica*, 41, 617–631.

Gui, H., R. Muller, and R. Vohra (2004): "Characterizing Dominant Strategy Mechanisms with Multidimensional Types," Working Paper, Northwestern University.

Jehiel, P., M. M. ter Vehn, and B. Moldovanu (2008): "Ex-Post Implementation and Preference Aggregation via Potentials," *Economic Theory*, 37, 469–490.

Lavi, R. (2007): *Algorithmic Game Theory*, Cambridge University Press, chap. Computationally-Efficient Approximate Mechanisms, 301–330, editors: Noam Nisan and Tim Roughgarden and Eva Tardos and Vijay Vazirani.

Lavi, R., A. Mualem, and N. Nisan (2003): "Towards a Characterization of Truthful Combinatorial Auctions," in *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS'03)*, IEEE Press.

——— (2009): "Two Simplified Proofs of Roberts' Theorem," *Social Choice and Welfare*, 32, 407–423.

Meyer-ter-Vehn, M. and B. Moldovanu (2003): "Ex Post Implementation with Interdependent Valuations," Mimeo.

Milnor, J. (1954): *Decision Processes*, Wiley, New York, chap. Games Against Nature, 44–59, ediors: R.M. Thrall, C.H. Coombs, and R.L. Davis.

Moulin, H. (1983): *The Strategy of Social Choice*, North-Holland.

Muller, R., A. Perea, and S. Wolf (2007): "Weak Monotonicity and Bayes-Nash Incentive Compatibility," *Games and Economics Behavior*, 61, 344–358.

Reny, P. (2001): "Arrow's Theorem and the Gibbard-Satterthwaite Theorem: A Unified Approach," *Economics Letters*, 70, 99–105.

Roberts, K. (1979): *The Characterization of Implementable Choice Rules*, North Holland Publishing, chap. Aggregation and Revelation of Preferences, 321–348, editor: J-J. Laffont.

Rochet, J. C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-linear Context," *Journal of Mathematical Economics*, 16, 191–200.

Saks, M. E. and L. Yu (2005): "Weak Monotonicity Suffices for Truthfulness on Convex Domains," in *Proceedings of $7^{th}$ ACM Conference on Electronic Commerce*, ACM Press, 286–293.

Trockel, W. (1992): "An Alternate Proof for the Linear Utility Representation Theorem," *Economic Theory*, 2, 298–302.

Vickrey, W. (1961): "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16, 8–37.

Vohra, R. V. (2008): "Paths, Cycles and Mechanism Design," Manuscript, Kellogg School of Management, Northwestern University.

——— (2011): *Mechanism Design: A Linear Programming Approach*, Cambridge University Press.

# Appendix A

Proof of Fact 1

*Proof*: Let $x, y \in X$ and $z = \alpha x + (1-\alpha)y$ for some $\alpha \in (0,1)$. We consider two possible cases.

Case 1: Suppose $X$ is closed in $D$. Assume for contradiction that $z \notin X$. Since $D$ is convex, $z \in D \setminus X$. Since $X$ is closed in $D$, the set $D \setminus X$ is open in $D$. Hence, $D \setminus X$ is open in $\mathbb{R}^n$. This means, there exists an $n$-dimensional open ball $B_\delta(z) = \{z' : \|z' - z\| < \delta\}$ of radius $\delta$ such that every $z' \in B_\delta(z)$ belongs to $D \setminus X$.

Now, consider an iterative procedure as follows. Let $l, h$ be two variables in $\mathbb{R}^n$. Initially, set $l = x$ and $h = y$. In every step,

- if $z$ is in the convex hull of $l$ and $\frac{l+h}{2}$ then set $h = \frac{l+h}{2}$,
- else set $l = \frac{l+h}{2}$.

If $\|l - h\| < 2\delta$, stop. Else, repeat the step.

Since $\|l - h\|$ strictly decreases in every step, the procedure will terminate. Moreover, $l$ and $h$ at the end of the procedures are two points in $X$. Hence, $\frac{l+h}{2}$ is in $X$ and lies in the ball $B_\delta(z)$. This is a contradiction [8].

[8] Essentially, the procedure generates a sequence of dyadic rational numbers. We know that the set of dyadic rational numbers are dense. Since $X$ is closed, we are done.

Case 2: Suppose $X$ is open in $D$. Then $X$ is open in $\mathbb{R}^n$. This implies that there exists an open ball $B_{\delta_x}(x)$ around $x$ of radius $\delta_x$ and an open ball $B_{\delta_y}(y)$ around $y$ of radius $\delta_y$ such that each of these balls are contained in $X$. Let $\delta = \min(\delta_x, \delta_y)$. Using the fact that for every $x' \in B_{\delta_x}(x)$ and every $y' \in B_{\delta_y}(y)$ we have $\frac{x'+y'}{2} \in X$, we get that every $x'' \in B_\delta(\frac{x+y}{2})$ lies in $X$. Now, we can repeat the procedure of Case 1 to conclude that $z \in X$. ■

Proof of Theorem 4

*Proof*: Note that $D$ is open from above (i.e., for every $x \in D$, there exists an $\epsilon \in \mathbb{R}^n_{++}$ such that $(x+\epsilon) \in D$) and a meet-semilattice (i.e., if $x, y \in D$, then $\min(x, y) \in D$). We can verify that Propositions 2 and 3 are true as long as $D$ is open from above and a meet-semilattice. Hence, by Proposition 2, $R^f$ is an ordering. By Proposition 3 and Lemma 5, $f$ satisfies weak Pareto, tr-invariance, and anonymity (we do not need continuity for this proof). Also, note that for any $x \in D$, any permutation of the elements of $x$ results in a vector in $D$.

Now, choose $x, y \in D$ such that $\sum_{i\in N} x_i = \sum_{i\in N} y_i$. By anonymity, we can rearrange $x$ and $y$ in non-decreasing order but mutually ranked the same way as $x$ and $y$. Considering successively, in these new vectors, each pair of corresponding components and subtracting from each the minimal one, we get again two new vectors which are ranked the same way as $x$ and $y$ by tr-invariance (note here that these two new vectors belong to $D = [0, H)^n$). Repeating these two operations at most $n$ times, we will reach two zero vectors (since $\sum_{i\in N} x_i = \sum_{i\in N} y_i$). Hence, $xI^f y$.

Next, we show that if $\sum_{i\in N} x_i > \sum_{i\in N} y_i$ then $xP^f y$. Let $\delta = \frac{1}{n}[\sum_{i\in N} x_i - \sum_{i\in N} y_i]$. Consider the vector $z$ defined as $z_i = y_i + \delta$ for all $i \in N$. By weak Pareto $zP^f y$. Further $\sum_{i\in N} x_i = \sum_{i\in N} z_i$. Hence, $xI^f z$. Hence, $xP^f y$.

By Lemma 2, for every $t \in T$, we have $f(t) \in C^f(t)$. Hence, $f(t)R^f a$ for all $a \in A$. Hence, $f$ is the efficient social choice function. ■