

Social Choice for Agents with General Utilities

Hongyao Ma

Harvard University
hma@seas.harvard.edu

Reshef Meir

Technion - Israel Inst. of Technology
reshefm@ie.technion.ac.il

David C. Parkes

Harvard University
parkes@eecs.harvard.edu

Abstract

The existence of truthful social choice mechanisms strongly depends on whether monetary transfers are allowed. Without payments there are no truthful, non-dictatorial mechanisms under mild requirements, whereas the VCG mechanism guarantees truthfulness along with welfare maximization when there are payments and utility is quasi-linear in money. In this paper we study mechanisms in which we can use payments but where agents have non quasi-linear utility functions. Our main result extends the Gibbard-Satterthwaite impossibility result by showing that, for two agents, the only truthful mechanism for at least three alternatives under general decreasing utilities remains dictatorial. We then show how to extend the VCG mechanism to work under a more general utility space than quasi-linear (the “parallel domain”) and show that the parallel domain is maximal—no mechanism with the VCG properties exists in any larger domain.

1 Introduction

Since its foundation, researchers interested in multi-agent AI have studied social choice mechanisms as a way to achieve coordinated decision-making. For example, social choice mechanisms were studied in the context of automated negotiation [Rosenschein and Zlotkin, 1994; Ephrati and Rosenschein, 1996] and for distributed, rational decision making [Sandholm, 1999].

A typical concern is that the mechanism be truthful, meaning that the optimal behavior for each agent is to make a truthful report about its private input (i.e., its preferences on outcomes.) When discussing truthful mechanisms, a sharp distinction is typically made based on whether the mechanism is allowed to use monetary transfers or not.

In the classical voting problem without money, the seminal Gibbard-Satterthwaite theorem [Gibbard, 1973; Satterthwaite, 1975] states that the only deterministic and truthful mechanisms for three candidates or more are dictatorial. This is true even if there are only two voters. On the other hand, the introduction of monetary transfers enables a value-maximizing, truthful mechanism known as the Clarke pivot rule, or VCG mechanism [Clarke, 1971; Groves, 1973;

Green and Laffont, 1979]. The mechanism collects cardinal utilities from agents, selects the alternative with the highest total bids and charges each agent the amount by which she affected the outcome. See Conitzer [2010] for a broader comparison of voting and auction mechanisms.

Due to its generality and the wide range of applications, the VCG mechanism has become one of the most common and powerful tools in mechanism design [Nisan, 2007]. One major drawback of VCG is that it relies very strongly on the assumption that agents’ utilities are *quasi-linear* (QL), i.e. of the form $u_a(z) = v_a - z$ for alternative a and payment z .¹

Non-QL utilities have gained much attention in recent years, both in AI and in economics. We provide some examples, demonstrating how these can arise in the context of group decisions and voting:

1. The use of contingent payments that depend on future decisions, such as a fine for agents who fail to use an item or show up for a meeting, transform QL utilities to ones that are convex in payment, since the amount of fine influences probability of follow through [Ma *et al.*, 2015].²
2. The time of a payment may depend on the selected alternative; e.g., deciding how and when to implement a project may affect the payment schedule. Together with temporal inconsistency, the utility from early alternatives will decrease more quickly as payment increases [Frederick *et al.*, 2002].
3. Agents may have hard or soft budget constraints in the context of automated bidding systems [Borgs *et al.*, 2005].

In this paper, we consider agents that have arbitrary utility functions (that are strictly decreasing with payment), and ask whether there exists a mechanism for aggregating preferences that is truthful (i.e. dominant-strategy incentive compatible), individually rational, unanimous, and never pays the agents. In other words, we ask if VCG or similar mechanisms can be extended to domains where utilities are non quasi-linear.

Contribution: Our main result is analogous to the Gibbard-Satterthwaite impossibility theorem for general non-quasi lin-

¹Indeed, for quasi-linear utility, and a rich enough value domain, any truthful mechanism has a structure that we can think of as “weighted VCG” [Roberts, 1979].

²Such fines are used for example to deter amateur soccer teams from failing to show up for scheduled matches (see rule 19, <http://www.footiefives.com/?q=node/394>).

ear domains (except it is restricted to two agents and unanimous rules) :

Theorem (Informal). *The only social choice mechanism for two voters with general decreasing utility functions that is truthful, unanimous, and never pays the agents, is dictatorial.*

If the designer also requires neutrality and anonymity, both of which are satisfied by the VCG mechanism, we show that such a mechanism exists if and only if the type space satisfies a *parallel domain* property, where for each agent, the gaps between the willingness to pay for different alternatives are constant.

Since the disruptive papers of Gibbard and Satterthwaite, researchers in AI and computational social choice have contributed a plethora of tools and techniques to mitigate and override their negative results, including computational complexity [Conitzer and Sandholm, 2002; Hemaspaandra *et al.*, 2007], uncertainty [Walsh, 2007; Conitzer *et al.*, 2011], iterative voting mechanisms [Meir *et al.*, 2010; Rabinovich *et al.*, 2015], commitments [Xia and Conitzer, 2010], automated mechanism design [Sandholm *et al.*, 2007; Jurca *et al.*, 2009] and more. Given that our main result is negative, we believe that it is worth exploring how these kinds of techniques can be applied to the more general social choice problem that we present in this paper.

2 Preliminaries

Let $N = \{1, 2, \dots, n\}$ be a set of *agents* and $A = \{a, b, \dots, m\}$ be a set of *alternatives*. We denote the set of all linear orders (permutations) over A by $\pi(A)$. For a vector $y = (y_1, y_2, \dots, y_n)$ we denote by y_{-i} the partial vector with all entries except y_i .

A *social choice mechanism* f accepts reports from agents as input, then outputs a single winner $a \in A$ and possibly payments that each agent should pay or receive. In classic voting, where monetary transfers are not allowed, it is assumed that the input is an ordinal *preference profile* $P = (\succ_i)_{i \in N}$, where $\succ_i \in \pi(A)$, and agent i prefers j to j' if $j \succ_i j'$. A social choice mechanism (a.k.a. a voting rule) is a function $f : \pi(A)^n \rightarrow A$ from profiles to a single winner.

f is *anonymous* if the outcome is independent of the names of the agents, and *neutral* if the outcome is independent of the names of the alternatives. f is *unanimous* if alternative j is selected as long as j is the favorite alternative of all agents.

Since agents are assumed to be self-interested, and their preferences are private information, they may report strategically in order to achieve a more preferred outcome. A mechanism is *dominant-strategy incentive compatible* (DSIC, sometimes called *strategyproof* or *truthful*), if no agent can gain by reporting false preferences.

f is a *dictatorship* if there exists an agent i s.t. her favorite alternative is always selected. The Gibbard-Satterthwaite Theorem [Gibbard, 1973; Satterthwaite, 1975] states that for any $m \geq 3, n \geq 2$, the only voting rules that are DSIC and unanimous are dictatorial.

2.1 Social Choice with Monetary Transfers

We next move to a more general framework, where mechanisms are allowed to make monetary transfers and the utility

of an agent may depend both on the selected alternative and her assigned payment. Denote $u_{i,j}(z)$ as the utility of agent i if alternative j is selected and she needs to pay z . We require that $u_{i,j}(z)$ is a continuous and strictly decreasing function of the payment z ,³ with $u_{i,j}(0) \geq 0$ and $u_{i,j}(\infty) \leq 0, \forall i, j$. We refer to $u_i = (u_{i,1}, \dots, u_{i,m})$ as agent i 's *type*, and to $u = (u_1, \dots, u_n)$ as a *type profile*.

Denote the utility of alternative j to agent i at zero payment as $v_{i,j} = u_{i,j}(0)$, which we call the “value.” As is common in social choice, $v_{i,j} \neq v_{i,j'}$ for all i for all j, j' . We thus retain notation $j \succ_i j'$ to denote that $v_{i,j} > v_{i,j'}$. Let $\tilde{j}_i \triangleq \arg \max_{j \in A} v_{i,j}$ and $\underline{j}_i \triangleq \arg \min_{j \in A} v_{i,j}$ be the most and least preferred alternative at zero payments for agent i .

Let U_0 be the set of all utility functions (satisfying our technical properties), and $\mathcal{U}_0 = U_0^n$ be the set of all profiles. Let $\mathcal{U} = U^n$ be the set of utility profiles that we work with. U is *quasi-linear* if $\forall u_i \in U, u_{i,j}(z) = v_{i,j} - z, \forall j \in A, \forall z \in \mathbb{R}$. Let \mathcal{U}_{QL} be the set of QL utility profiles where agent values are non-negative.

Thus a *social choice mechanism* on type space \mathcal{U} is composed of a *choice rule* $x : \mathcal{U} \rightarrow A$ and a *payment rule* $t = (t_1(u), \dots, t_n(u)) : \mathcal{U} \rightarrow \mathbb{R}^n$. Thus if the choice made is $x(u) = j$ then each agent i gains $u_{i,j}(t_i(u))$.

The definitions of Anonymity and Neutrality naturally extend to mechanisms with payments. A mechanism (x, t) is DSIC if for any type $u_i \in U$ of any agent $i \in N$ and any reported profile from other agents $\hat{u}_{-i} \in U^{n-1}$, agent i cannot gain by misreporting any false type $\hat{u}_i \in U$:

$$u_{i,x(u_i, \hat{u}_{-i})}(t_i(u_i, \hat{u}_{-i})) \geq u_{i,x(\hat{u}_i, \hat{u}_{-i})}(t_i(\hat{u}_i, \hat{u}_{-i}))$$

A mechanism is *Individually Rational* (IR) if $\forall i \in N$,

$$u_{i,x(u_i, \hat{u}_{-i})}(t_i(u_i, \hat{u}_{-i})) \geq 0, \forall u_i \in U, \forall \hat{u}_{-i} \in U^{n-1}$$

A mechanism satisfies *unanimity* if when all agents agree on which alternative is the best at zero-payment, this alternative is selected and all payments are zero. We are especially interested in mechanisms that hold following set of conditions:

- A1. Dominant-strategy incentive compatible (DSIC)
- A2. No payments from the mechanism to agents
- A3. Deterministic (unless breaking ties)
- A4. Unanimous
- A5. Anonymous
- A6. Neutral
- A7. Individually Rational (IR)

Requirements A1, A2, and A7 are standard in mechanism design, whereas A3–A6 are typically required in social choice.⁴

2.2 Agent-Independent Prices

The structure of DSIC mechanisms with payments is well known (see [Nisan, 2007] for an overview). In particular, every DSIC mechanism must be *agent independent* and *agent maximizing*.

³In voting without money $u_{i,j}(z)$ is a *constant*.

⁴We require the mechanism to be deterministic unless there are multiple alternatives that are simultaneously agent-maximizing for every agent. Only in this case can the mechanism break ties at random. Allowing randomization among outcomes among which agents may have strict preferences, and for the special case of two alternatives, a majority vote without payments and with random tie-breaking satisfies all of the requirements A1-A7.

Lemma 1 (Agent Independence). For any DSIC social choice mechanism (x, t) , $\forall i \in N$, $\forall u_i, \forall u'_i, \forall u_{-i}$, if $x(u_i, u_{-i}) = x(u'_i, u_{-i})$, then $t_i(u_i, u_{-i}) = t_i(u'_i, u_{-i})$.

Intuitively, fixing the report u_{-i} of the other agents, an agent i cannot affect her own payment without affecting the chosen alternative. Otherwise, an agent with type u_i has an incentive to report the type u'_i or vice versa.

Given u_{-i} , if there exists $u_i \in U$ s.t. $x(u_i, u_{-i}) = j$, then let the *agent-independent prices* be the payment i pays when j is selected: $t_{i,j}(u_{-i}) \triangleq t_i(u_i, u_{-i})$, which depends only on u_{-i} . Otherwise, let $t_{i,j}(u_{-i}) \triangleq \infty$.

Lemma 2 (Agent Maximization). For any DSIC social choice mechanism (x, t) , $\forall u \in \mathcal{U}$, there must exist $j^* \in A$ s.t. $x(u) = j^*$ and $j^* \in \arg \max_{j \in A} u_{i,j}(t_{i,j}(u_{-i}))$, $\forall i \in N$, i.e. the selected alternative must be simultaneously agent maximizing under the agent-independent prices for all agents.

It is easy to see that if agent-maximization does not hold there would be useful deviations. Further, any mechanism that meets both requirements is DSIC: an agent cannot affect her price for any particular alternative, and at the given prices (fixed by reports of others), she has no incentive to report in a way that results in a different alternative being selected. Therefore, a social choice mechanism is DSIC if and only if it can be stated in the following form:

- Compute agent-independent prices $t_{i,j}(u_{-i})$ for each agent i and each alternative j .
- Choose an alternative $j^* \in \arg \max_{j \in A} u_{i,j}(t_{i,j})$, and charge each agent $t_i(u) = t_{i,j^*}(u_{-i})$.

As an example, a simple mechanism that always chooses alternative a and collects no payment sets $t_{i,a} = 0$ and $t_{i,j} = \infty$, for all $j \neq a$, for all $i \in N$. It is easy to see that these payments are agent-independent and choosing a is the agent-maximizing allocation for all agents. For simplicity of notation, we drop the argument and denote $t_{i,j} \triangleq t_{i,j}(u_{-i})$ when there is no ambiguity here onwards.

VCG mechanism. The Vickrey-Clarke-Groves (VCG) mechanism is defined for QL utilities and computes the value-maximizing outcome, collecting from each agent the negative externality that she imposes on the rest of the agents. With QL utilities, all the information on an agent's type is conveyed in her values $v_{i,j}$. Thus, in the context of social-choice, VCG asks for a bid from each agent for each alternative, selects the alternative that has the maximum total bid value, and charges each agent the amount by which she is "pivotal" (meaning the amount of her bid that is important in causing this alternative to be selected.)

Mechanism 1: VCG Mechanism (QL Utilities)

- Input: Bids $b_{i,j}$ for any $i \in N$ and $j \in A$
Set $x(b) \triangleq j^* = \arg \max_{j \in A} \sum_{i \in N} b_{i,j}$
// allocation rule
Set $t_i(b) = \max_j \sum_{i' \neq i} b_{i',j} - \sum_{i' \neq i} b_{i',j^*}$
// payment rule
-

In the case of ties, i.e. two alternatives with the same total bid, we can break ties randomly. The mechanism is deterministic in cases other than random tie-breaking.

For QL utilities, the VCG mechanism satisfies all properties A1–A7 (see [Nisan, 2007]). In particular, it is a dominant strategy for agents to bid their true values on each alternative $b_{i,j}^* = v_{i,j}$. Observe that for any v_{-i} , there exists types v_i with very large $v_{i,j}$ such that alternative j would be selected. Denote $j_{-i}^* = \arg \max_{j \in A} \sum_{i' \neq i} v_{i',j}$. For each alternative $j \in A$, the agent-independent prices are of the form $t_{i,j}(v_{-i}) = \sum_{i' \neq i} v_{i',j_{-i}^*} - \sum_{i' \neq i} v_{i',j}$. It is easy to verify that the welfare-maximizing alternative j^* is agent-maximizing for all agents.

Dictatorship. A *dictator mechanism* is defined by identifying a particular agent $i^* \in N$ and defining x to select alternative j to maximize $v_{i^*,j}$, with all payments set to 0 (i.e. $t_{i^*}(u) \equiv 0$). A dictatorship trivially satisfies all of our required properties except anonymity. For the dictator, $t_{i^*}(u)$ is always zero thus $t_{i^*,j}(u_{-i^*}) = 0$ always holds for all j regardless of the reports from the others u_{-i^*} . Let $j^* = \bar{j}_{i^*}$ be the dictator's favorite alternative. In order for all other agents to always prefer alternative j^* at their corresponding prices, we must have $t_{i,j^*} = 0$ and $t_{i,j} = \infty$ for all $j \neq j^*$, for all $i \neq i^*$. These payments are agent-independent, and alternative j^* is agent-maximizing for all agents.

3 Impossibility under General Utilities

We show that for general non-QL utility space \mathcal{U}_0 with two agents and at least three alternatives, the only mechanism satisfying A1–A4 must be a dictatorship. We focus w.l.o.g. on the true types as the inputs of the mechanism since we study DSIC mechanisms.

Let $j_i^{(k)}$ be the k th favorite alternative for agent i at zero payment, i.e. $v_{i,j_i^{(1)}} > v_{i,j_i^{(2)}} > \dots > v_{i,j_i^{(m)}}$. $t_{-i,j_i^{(k)}}(u_i)$ denotes the price the other agent $-i$ faces for the alternative that i ranks at place k , which is a function of only agent i 's type u_i due to agent independence. For a particular social choice mechanism, we define two subsets of the type space $U_i^I, U_i^{II} \subseteq U_0$ as:

- $U_i^I = \{u_i \in U_0 \mid t_{-i,j_i^{(2)}}(u_i) = 0\}$
- $U_i^{II} = \{u_i \in U_0 \mid t_{-i,j_i^{(k)}}(u_i) = \infty, \forall k \geq 3\}$

U_i^I (utilities of the first type) is the collection of i 's types that set a zero price for the other agent on i 's second favourite alternative. U_i^{II} (utilities of the second type) is the collection of i 's types s.t. all but the first two favourite alternatives for i are never selected for any u_{-i} . We first prove that any agent type must belong to at least one of these two collections.

Lemma 3. In any mechanism satisfying A1–A4 for the non-QL utility space \mathcal{U}_0 with two agents, then one of $u_i \in U_i^I$ and $u_i \in U_i^{II}$ must hold for any agent i and any $u_i \in U_0$.

Proof. Assume otherwise, there exists an agent i and a type $u_i \in U_0$ s.t. $u_i \notin U_i^I$ and $u_i \notin U_i^{II}$. W.l.o.g. we assume that $\exists u_1 \in U_0$ s.t. $a \succ_1 b \succ_1 c \succ_1 \dots \succ_1 m, t_{2,b}(u_1) > 0$

and $t_{2,j}(u_1) < \infty$ for some $j \neq a, b$. Consider for now that $j = c$, and the same argument can be repeated for all $j \neq a, b$.

With all possible types in U_0 , we can find u_2 of agent 2 as shown in Figure 1 s.t. $b \succ_2 c \succ_2 a$ and $u_{2,c}(t_{2,c}) > u_{2,b}(t_{2,b}) > v_{2,j}$, $\forall j \neq b, c$, assuming $t_{2,b} > 0$ and $t_{2,c} < \infty$.

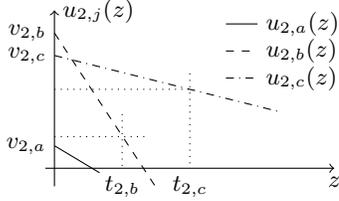


Figure 1: Preference u_2 of agent 2 for the proof of Lemma 3.

We know from unanimity that $t_{1,b} = 0$ must hold, since if agent 1 reports b as her favorite alternative, b must be selected and her payment must be zero. Therefore we have $\forall j \neq a, b$, $u_{1,j}(t_{1,j}) \leq v_{1,j} < v_{1,b} = u_{1,b}(t_{1,b})$, and the only possible agent-maximizing alternatives for agent 1 are a or b .

However, it is easy to see that the agent-maximizing alternative for agent 2 is c , thus there does not exist an alternative that is agent-maximizing for both agents. Contradiction. \square

The placement of agent i 's type into U_1^{II} or U_2^{II} depends only on the mechanism and u_i , and not on the report of the other agent. Say that an agent i is a *dictator for alternative j* , if $\bar{j}_i = j \Rightarrow x(u) = j$, and $t_{i,j}(u_{-i}) = 0$ for all $u_{-i} \in U_0$. Next, we show that if there exists $u_i \in U_i^I$, the other agent must be a dictator for $j_i^{(2)}$, the second favorite alternative of agent i .

Lemma 4. In any mechanism satisfying A1–A4 for the non-QL utility space \mathcal{U}_0 with two agents, if there exists a type $u_i \in U_i^I$ with $j = j_i^{(2)}$, then agent $-i$ is a dictator for j .

Proof. Assume w.l.o.g. that there exists a type $u_2 \in U_2^I$ s.t. $b \succ_2 a \succ_2 j$ for all $j \neq a, b$. We show that agent 1 is a dictator for alternative a . We proceed in steps:

Claim 1: For $u_1 \in U_0$ s.t. $\bar{j}_1 = a$ and $j_1^{(2)} = b$, we have $u_1 \in U_1^{II}$. Assume otherwise, so that both u_1 and u_2 belong to U_1^I and $t_{1,a} = t_{2,b} = 0$. We know from unanimity that $t_{2,a} = t_{1,b} = 0$. Under these prices, $u_{1,a}(t_{1,a}) = v_{1,a} > v_{1,b} = u_{1,b}(t_{1,b})$ and $u_{2,a}(t_{2,a}) = v_{2,a} < v_{2,b} = u_{2,b}(t_{2,b})$, and the agent-maximizing alternatives are different. A contradiction.

Claim 2: For $u_2 \in U_0$ s.t. $c \succ_2 a \succ_2 j$ for $j \neq a, c$ and $c \neq b$, we have $u_2 \in U_2^I$. Assume otherwise, $u_2 \notin U_2^I$. This implies $t_{1,j_2^{(2)}}(u_2) = t_{1,a}(u_2) \neq 0$ thus $t_{1,a}(u_2) > 0$ since no payment is made to the agents. We can find a type $u_1 \in U_0$ as shown in Figure 2 s.t. $a \succ_1 b \succ_1 j$ for all $j \neq a, b$, but $u_{1,a}(t_{1,a}) < v_{1,c}$. From unanimity, we have $t_{2,a} = 0$, and thus the agent maximizing alternative for agent 2 must be a or c . Further, we know from Claim 1 that $u_1 \in U_1^{II}$ thus $t_{2,c} = \infty$ so alternative a must be selected. However, again

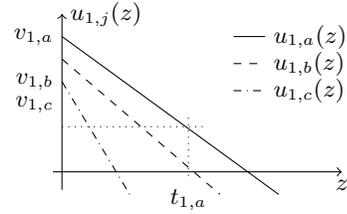


Figure 2: Preference u_1 of agent 1 for Claim 2 of Lemma 4.

from unanimity we know $t_{1,c} = 0$ thus $u_{1,a}(t_{1,a}) < v_{1,c} = u_{1,c}(t_{1,c})$. A contradiction.

Claim 3: $\forall u_1 \in U_0$ s.t. $\bar{j}_1 = a$, we have $u_1 \in U_1^{II}$. With some $u_2 \in U_2^I$ s.t. $b \succ_2 a \succ_2 j$ for $j \neq a, b$, we showed in Claim 1 that for u_1 s.t. $\bar{j}_1 = a$ and $j_1^{(2)} = b$, $u_1 \in U_1^{II}$. We know from Claim 2 that for all $c \neq a, b$, there exists $u_2 \in U_2^I$ s.t. $c \succ_2 a \succ_2 j$ for $j \neq a, c$. Repeating the same argument as in Claim 1, we know that for all $c \neq a, b$, $\bar{j}_1 = a$ and $j_1^{(2)} = c$ implies $u_1 \in U_1^{II}$. Combining this with Claim 1, we know $\forall u_1 \in U_0$ s.t. $\bar{j}_1 = a$, $u_1 \in U_1^{II}$.

Claim 4: For $u_2 \in U_0$ s.t. $j_2^{(2)} = a$, we have $u_2 \in U_2^I$. In Claim 2, we showed that for u_2 s.t. $\bar{j}_2 = c \neq b$ and $j_2^{(2)} = a$, $u_2 \in U_2^I$. What is left to show is that $u_2 \in U_2^I$ for u_2 s.t. $\bar{j}_2 = b$ and $j_2^{(2)} = a$. This follows by repeating the argument in Claim 2, observing that Claim 1 is generalized to u_1 with $\bar{j}_1 = a$ and any $j_1^{(2)}$.

Claim 5: For all $u_1 \in U_0$ s.t. $\bar{j}_1 = a$, we have $t_{2,j}(u_1) = \infty$ for all $j \neq a$. We know from Claim 3 that $t_{2,j} = \infty$ for all $j \neq a, j_1^{(2)}$, and left to show is $t_{2,j_1^{(2)}} = \infty$. Denote $j_1^{(2)} = b$ w.l.o.g. and assume for contradiction that $t_{2,b} < \infty$. There exists u_2 of agent 2 s.t. $b \succ_2 a \succ_2 j$ for all $j \neq a, b$ and $u_{2,b}(t_{2,b}) > v_{2,a}$. We know from Claim 4 that $t_{1,a} = 0$, thus a is agent-maximizing for agent 1 and must be selected. However, agent 2 strictly prefers b to a at the current prices, thus this is a contradiction and $t_{2,b}(u_1) = \infty$ must hold.

Claim 6: $\forall u_2 \in U_0$, we have $t_{1,a}(u_2) = 0$. Assume otherwise, such that there exists u_2 s.t. $t_{1,a}(u_2) > 0$. We can find a type u_1 of agent 1 such that $a \succ_1 j$ and $u_{1,a}(t_{1,a}) < v_{1,j}$ for all $j \neq a$. From unanimity, we know $t_{1,\bar{j}_2} = 0$ thus $u_{1,a}(t_{1,a}) < v_{1,\bar{j}_2} = u_{1,\bar{j}_2}(t_{1,\bar{j}_2})$. This shows that a is not the agent-maximizing alternative for agent 1, however from Claim 5 alternative a must be selected. A contradiction.

Thus we know that agent 1 never pays for alternative a , and a is selected as long as it is agent 1's most preferred alternative. This implies that agent 1 is a dictator for a . \square

Theorem 1. With two agents and at least three alternatives, the only social choice mechanism for the non-QL utility space \mathcal{U}_0 satisfying A1–A4 must be a dictatorship.

Proof. We proceed in steps:

Step 1: It is not possible that $U_1^I = U_2^I = U_0$. Otherwise, consider any profile where $a \succ_1 b \succ_1 \dots; b \succ_2 a \succ_2 \dots$,

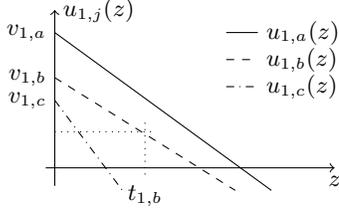


Figure 3: Preference u_1 of agent 1 for Step 2 of Theorem 1.

we get a contradiction to agent maximizing with the same arguments as in Claim 1 of the proof of Lemma 4.

Step 2: It is not possible that $U_1^{II} = U_2^{II} = U_0$. This is obvious if the number of alternatives is at least 4 since there would be type profiles where all alternatives are eliminated. Assume there are three alternatives and $U_1^{II} = U_2^{II} = U_0$.

We know from Claim 1 that there must be some type of some agent that is not in U^I ; thus, we assume w.l.o.g. $\exists u_2 \notin U_2^I$ and we name the alternatives s.t. $c \succ_2 b \succ_2 a$.

We know from the assumptions that $t_{1,b} > 0$. Consider the type u_1 as shown in Figure 3 (we do not need to specify u_2 exactly), i.e. for which $u_{1,b}(t_{1,b}) < v_{1,c}$. Since $u_1 \in U_1^{II}$, $u_2 \in U_2^{II}$, the only alternative that has not been rejected is b . However, $t_{1,c} = 0$ by unanimity thus $u_{1,c}(t_{1,c}) = v_{1,c} > u_{1,b}(t_{1,b})$. This contradicts agent-maximization, thus $U_1^{II} = U_2^{II} = U_0$ cannot hold.

Step 3: By Step 2, the premise of Lemma 4 holds for some $i \in \{1, 2\}$ and some alternative. Thus w.l.o.g. agent 1 is a dictator for a , and $t_{1,a} = 0$ for all $u_2 \in U$.

Step 4: For any alternative $j \neq a$, consider the type u_2 s.t. j is ranked second and a is ranked last: $j_2^{(2)} = j$ and $j_2^{(m)} = a$. Since $t_{1,a}(u_2) = 0 < \infty$ must hold, $u_2 \notin U_2^{II}$ thus $u_2 \in U_2^I$ by Lemma 3. Again by Lemma 4 we know that agent 1 is the dictator for alternative j .

Therefore agent 1 is the dictator for all alternatives. \square

We know that the VCG mechanism meets all our requirements A1–A7 for QL utilities. Theorem 1 shows that once we allow arbitrary decreasing utility functions, then even minimal requirements A1–A4 lead to an impossibility result. A natural question is whether a positive result is available for a smaller set of non-QL utilities. Our impossibility result holds as long as we allow utility functions that are sufficiently “shallow.” For example, all decreasing linear functions. In the remainder of the paper we explore the implications of more restricted utilities.

4 From Quasi-Linear to Parallel Utilities

We now define the concept of a *parallel domain* for utilities. A type profile in a parallel domain is one in which, for each agent, the horizontal distances between the utility curves of different alternatives are constant, whenever the utility curves are above the value of the least preferred alternative.

We present a direct mechanism that satisfies A1–A7 for parallel domains, and prove an impossibility result for type spaces that are not parallel.

Definition 1 (Parallel Domain). A type space $\mathcal{U} \subseteq \mathcal{U}_0$ is a parallel domain if for all $u \in \mathcal{U}$, for all $i \in N$, for all $j, j' \in A$ s.t. $v_{i,j} > v_{i,j'}$,

$$u_{i,j}(z + u_{i,j}^{-1}(v_{i,j'})) = u_{i,j'}(z), \forall 0 \leq z \leq u_{i,j}^{-1}(v_{i,j_i}).$$

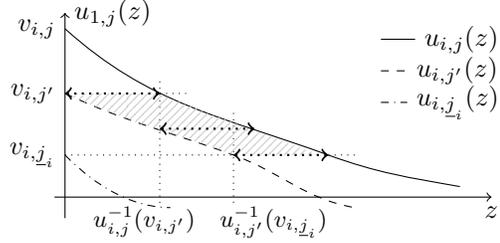


Figure 4: An example utility type in the parallel domain. All horizontal sections of the shaded area (e.g. the thick arrows) are of the same length, i.e. utility curves are horizontal shifts of each other within the range.

In other words, $\forall w, \in [v_{i,j_i}, v_{i,j'}]$, $u_{i,j}^{-1}(w) - u_{i,j'}^{-1}(w)$ is a constant that does not depend on w , i.e. the additional amount an agent is willing to pay for j' over j does not depend on how much the agent is charged for j . There is no requirement on the shape of the curves $u_{i,j}$ below v_{i,j_i} . Denote $\mathcal{U}_{\parallel} \subseteq \mathcal{U}_0$ as the largest parallel domain.

The QL space \mathcal{U}_{QL} is a parallel domain. A type space in which all first $m - 1$ utility curves of each agent are parallel (but not necessarily parallel to the utility curves of other agents) is also a parallel domain, which we call “linear parallel.” For these two spaces, the vertical distances between the utility curves also stay the same, but this does not hold for general parallel domains. Finally, when there are only two alternatives, \mathcal{U}_0 is trivially a parallel domain.

We now introduce a direct mechanism that satisfies A1–A7 for any number of agents and any number of alternatives if the type space is a parallel domain: $\mathcal{U} \subseteq \mathcal{U}_{\parallel}$. Given a type profile $u \in \mathcal{U}$, denote the “maximum willingness to pay” $p_{i,j}$ as the amount at which agent i is indifferent between getting alternative j and getting the worst alternative for free: $p_{i,j} = u_{i,j}^{-1}(v_{i,j_i}) = u_{i,j}^{-1}(\min_{j' \in A} v_{i,j'})$.

Mechanism 2: Direct mechanism for parallel domains

Input: $\hat{u} = (\hat{u}_1, \dots, \hat{u}_n)$ for $i \in N, j \in A$ do
 [Set $p_{i,j} = \hat{u}_{i,j}^{-1}(\hat{v}_{i,j_i}) = \hat{u}_{i,j}^{-1}(\min_{j' \in A} \hat{v}_{i,j'})$
 Set $x(\hat{u}) \triangleq j^* = \arg \max_{j \in A} \sum_{i \in N} p_{i,j}$
 // allocation rule
 Set $t_i(\hat{u}) = \max_j \sum_{i' \neq i} p_{i',j} - \sum_{i' \neq i} p_{i',j^*}$
 // payment rule

Theorem 2. *The direct mechanism satisfies A1–A7 for parallel domains, for any number of voters and alternatives.*

Proof. A2, A3 and A5–A7 are easy to see. For A4, the willingness to pay is aligned with values at zero payment for par-

allel domains and $\arg \max p_{i,j} = \bar{j}_i$. If $\exists j^* = \bar{j}_i$ for all i , $t_{i,j^*} = 0$ for all i and the mechanism satisfies unanimity.

What is left to check is DSIC. We can check that the prices are of the form $t_{i,j}(\hat{u}_{-i}) = \max_{j'} \sum_{i' \neq i} p_{i',j'} - \sum_{i' \neq i} p_{i',j}$ and are agent independent by construction, thus we only need to show $x(\hat{u}) = j^* = \arg \max_{j \in A} \sum_{i \in N} p_{i,j}$ is simultaneously agent maximizing for every agent. Denote $j_{-i}^* = \arg \max_{j \in A} \sum_{i' \neq i} p_{i',j}$. Observe that for parallel domains:

$$(i) p_{i,\bar{j}_i} = u_{i,\bar{j}_i}^{-1}(v_{i,\bar{j}_i}) = 0$$

$$(ii) \forall j, j' \text{ s.t. } j \succ_i j', \text{ we have } p_{i,j} - p_{i,j'} = u_{i,j}^{-1}(v_{i,j'})$$

Also, we know from the payment rule that $t_{i,j^*} \leq p_{i,j^*}$. We need to discuss the following cases:

Case 1: $j^* = \bar{j}_i$ i.e. the worse alternative is selected. Since $p_{i,j^*} = p_{i,\bar{j}_i} = 0$, we must have $t_{i,j^*} = 0$ since $t_{i,j^*} \leq p_{i,j^*}$. Also from $p_{i,j^*} = 0$, $j^* = \bar{j}_i$ must hold. For any $j \neq j^*$, the price agent i faces $t_{i,j} = \sum_{i' \neq i} p_{i',j_{-i}^*} - \sum_{i' \neq i} p_{i',j} = \sum_{i'} p_{i',j^*} - p_{i,j^*} - \sum_{i'} p_{i',j} + p_{i,j} = (\sum_{i'} p_{i',j^*} - \sum_{i'} p_{i',j}) + p_{i,j} \geq p_{i,j}$. This implies $u_{i,j^*}(t_{i,j^*}) = u_{i,\bar{j}_i}(0) = v_{i,\bar{j}_i} = u_{i,\bar{j}_i}(p_{i,j}) \geq u_{i,j}(t_{i,j})$, and j^* is the agent maximizing alternative.

Case 2: $j^* \neq \bar{j}_i$. First, we know from $t_{i,j^*} \leq p_{i,j^*}$ that $u_{i,j^*}(t_{i,j^*}) \geq u_{i,j^*}(p_{i,j^*}) = v_{i,\bar{j}_i} = u_{i,\bar{j}_i}(0) \geq u_{i,\bar{j}_i}(t_{i,\bar{j}_i})$. What is left to show is that for $j \neq \bar{j}_i$, $u_{i,j^*}(t_{i,j^*}) \geq u_{i,j}(t_{i,j})$. Observe:

$$\begin{aligned} t_{i,j^*} - t_{i,j} &= \sum_{i' \neq i} p_{i',j_{-i}^*} - \sum_{i' \neq i} p_{i',j^*} - (\sum_{i' \neq i} p_{i',j_{-i}^*} - \sum_{i' \neq i} p_{i',j}) \\ &= - \sum_{i'} p_{i',j^*} + p_{i,j^*} + \sum_{i'} p_{i',j} - p_{i,j} \leq p_{i,j^*} - p_{i,j} \end{aligned}$$

We consider the following two cases:

Case 2.1 $j^* \prec_i j$. We know from (ii) that $p_{i,j} - p_{i,j^*} = u_{i,j}^{-1}(v_{i,j^*})$, thus $t_{i,j^*} - t_{i,j} \leq p_{i,j^*} - p_{i,j} = -u_{i,j}^{-1}(v_{i,j^*}) \Rightarrow t_{i,j} \geq t_{i,j^*} + u_{i,j}^{-1}(v_{i,j^*})$. Since $t_{i,j^*} \leq p_{i,j^*} = u_{i,\bar{j}_i}^{-1}(v_{i,\bar{j}_i})$, we know from the property of parallel domain that $u_{i,j}(t_{i,j}) \leq u_{i,j}(t_{i,j^*} + u_{i,j}^{-1}(v_{i,j^*})) = u_{i,j^*}(t_{i,j^*})$.

Case 2.2 $j^* \succ_i j$. We know that $p_{i,j^*} - p_{i,j} = u_{i,j^*}^{-1}(v_{i,j})$ therefore $t_{i,j^*} \leq u_{i,j^*}^{-1}(v_{i,j}) + t_{i,j}$. If $t_{i,j} > p_{i,j}$, we know that $u_{i,j}(t_{i,j}) < v_{i,j} = t_{i,j^*}(p_{i,j^*}) \leq u_{i,j^*}(t_{i,j^*})$ since $t_{i,j^*} \leq p_{i,j^*}$. If $t_{i,j} \leq p_{i,j}$, we know from the property of parallel domain that $u_{i,j^*}(t_{i,j^*}) \geq u_{i,j^*}(u_{i,j^*}^{-1}(v_{i,j}) + t_{i,j}) = u_{i,j}(t_{i,j})$.

Combining all above cases, we have showed that j^* is agent-maximizing for all agents. This completes the proof. \square

Remark: It is evident from the description of the mechanism that there is an equivalent indirect mechanism, where each agent i only reports a ‘‘bid’’ $p_{i,j}$ for each $j \in A$.

4.1 Parallel Domains are Tight

We present the characterization of agent-independent prices for two agents in any mechanism under A1–A7, and show that the parallel domain is the largest type space in which such a mechanism exists. The details of the proofs are omitted due to the space limit.

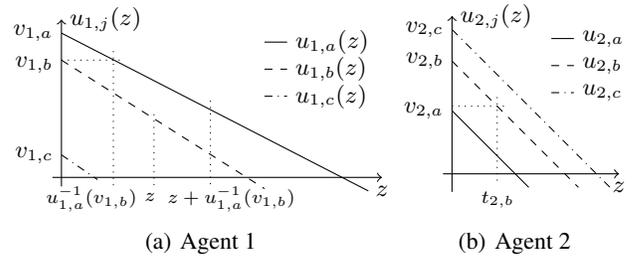


Figure 5: Preference profile the proof of Theorem 3.

Lemma 5 (Necessary Conditions for Agent-independent Prices). Consider two agents with types u_1 and u_2 . Under any social choice mechanism satisfying A1–A7, the agent independent prices for the other agent $-i$ determined by agent i ’s report $t_{-i,j}(u_i)$ must satisfy

$$t_{-i,j} = u_{i,\bar{j}_i}^{-1}(v_{i,j}), \quad (1)$$

i.e. for each alternative j , the other agent pays the amount that agent i needs to pay for \bar{j}_i such that agent i is indifferent between getting \bar{j}_i at this amount, and getting j for free.

Theorem 3. With two agents and at least three alternatives, under any type space \mathcal{U} s.t. $\mathcal{U}_{QL} \subset \mathcal{U}$ and $\mathcal{U} \not\subset \mathcal{U}_{\parallel}$, there is no social choice mechanism satisfying assumptions A1–A7.

Proof sketch for $m = 3$. Assume w.l.o.g that there exists a type u_1 of agent 1 s.t. $a \succ_1 b \succ_1 c$ that is non-parallel. There must exist some $z \leq u_{1,b}^{-1}(v_{1,c})$ s.t. $u_{1,b}(z) \neq u_{1,a}(z + u_{1,a}^{-1}(v_{1,b}))$, as shown in Figure 5(a). Suppose $u_{2,b}$ is more steep, as in the figure. We construct agent 2 with QL utilities $c \succ_2 b \succ_2 a$, s.t. agent 2 selects b but just barely. Thus the price she sets on a , due to the milder slope, results in agent 1 selecting a , in contradiction to agent-maximizing. \square

5 Conclusions

This paper has introduced the study of general social choice with non-QL utilities. We showed that without additional assumptions on the shape of agents’ utility functions, any truthful social choice mechanism for three alternatives or more (under mild requirements) is dictatorial. While our proof only applies for two voters, we believe that the result holds for any number of voters, and leave it as an open question.

We have also provided a tight answer to the question of how restricted utility functions must be in order for VCG-like mechanisms to work, showing that the QL domain can be somewhat extended to parallel domain, but no further.

Because utilities are likely to go beyond the parallel domain in most realistic applications, we argue that future work should relax the DSIC requirement, and see whether weaker incentive guarantees can be attained for example by factoring considerations of computational intractability on agent reasoning, or through behavioral models [Meir *et al.*, 2014].

Acknowledgments

This work is supported in part by a grant from the FLI, from the SEAS TomKat fund, and from NSF grant CCF-1301976.

References

- [Borgs *et al.*, 2005] Christian Borgs, Jennifer Chayes, Nicole Immorlica, Mohammad Mahdian, and Amin Saberi. Multi-unit auctions with budget-constrained bidders. In *ACM-EC'05*, pages 44–51, 2005.
- [Clarke, 1971] Edward H Clarke. Multipart pricing of public goods. *Public choice*, 11(1):17–33, 1971.
- [Conitzer and Sandholm, 2002] Vincent Conitzer and Tuomas Sandholm. Complexity of manipulating elections with few candidates. In *AAAI'02*, pages 314–319, 2002.
- [Conitzer *et al.*, 2011] Vincent Conitzer, Toby Walsh, and Lirong Xia. Dominating manipulations in voting with partial information. In *AAAI'11*, pages 638–643, 2011.
- [Conitzer, 2010] Vincent Conitzer. Comparing multiagent systems research in combinatorial auctions and voting. *Annals of Mathematics and Artificial Intelligence*, 58(3-4):239–259, 2010.
- [Ephrati and Rosenschein, 1996] Eithan Ephrati and Jeffrey S Rosenschein. Deriving consensus in multiagent systems. *Artificial Intelligence*, 87(1):21–74, 1996.
- [Frederick *et al.*, 2002] Shane Frederick, George Loewenstein, and T O'donoghue. Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2):351–401, 2002.
- [Gibbard, 1973] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica: journal of the Econometric Society*, pages 587–601, 1973.
- [Green and Laffont, 1979] Jerry R Green and Jean-Jacques Laffont. *Incentives in public decision making*. North-Holland, 1979.
- [Groves, 1973] Theodore Groves. Incentives in teams. *Econometrica: Journal of the Econometric Society*, pages 617–631, 1973.
- [Hemaspaandra *et al.*, 2007] Edith Hemaspaandra, Lane A Hemaspaandra, and Jörg Rothe. Anyone but him: The complexity of precluding an alternative. *Artificial Intelligence*, 171(5):255–285, 2007.
- [Jurca *et al.*, 2009] Radu Jurca, Boi Faltings, et al. Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research*, 34(1):209, 2009.
- [Ma *et al.*, 2015] Hongyao Ma, Reshef Meir, David C. Parkes, and James Zou. Are you going to do that? contingent-payment mechanisms to improve coordination. The Third Conference on Auctions, Market Mechanisms and Their Applications (AMMA'15), 2015.
- [Meir *et al.*, 2010] Reshef Meir, Maria Polukarov, Jeffrey S. Rosenschein, and Nick Jennings. Convergence to equilibria of plurality voting. In *AAAI'10*, pages 823–828, 2010.
- [Meir *et al.*, 2014] Reshef Meir, Omer Lev, and Jeffrey S. Rosenschein. A local-dominance theory of voting equilibria. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC '14, pages 313–330, 2014.
- [Nisan, 2007] Noam Nisan. Introduction to mechanism design (for computer scientists). In Noam Nisan et al., editor, *Algorithmic game theory*. Cambridge University Press, 2007.
- [Rabinovich *et al.*, 2015] Zinovi Rabinovich, Svetlana Obraztsova, Omer Lev, Evangelos Markakis, and Jeffrey S. Rosenschein. Analysis of equilibria in iterative voting schemes. In *AAAI'15*, pages 1007–1013, 2015.
- [Roberts, 1979] Kevin Roberts. The characterization of implementable choice rules. *Aggregation and revelation of preferences*, 12(2):321–348, 1979.
- [Rosenchein and Zlotkin, 1994] Jeffrey S Rosenschein and Gilad Zlotkin. *Rules of encounter: designing conventions for automated negotiation among computers*. MIT press, 1994.
- [Sandholm *et al.*, 2007] Tuomas Sandholm, Vincent Conitzer, and Craig Boutilier. Automated design of multistage mechanisms. In *IJCAI'07*, pages 1500–1506, 2007.
- [Sandholm, 1999] T. W. Sandholm. Distributed rational decision making. In G. Weiss, editor, *Multiagent Systems: A Modern Introduction to Distributed Artificial Intelligence*, pages 201–258. MIT Press, 1999.
- [Satterthwaite, 1975] Mark Allen Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217, 1975.
- [Walsh, 2007] Toby Walsh. Uncertainty in preference elicitation and aggregation. In *AAAI'07*, pages 3–8, 2007.
- [Xia and Conitzer, 2010] Lirong Xia and Vincent Conitzer. Stackelberg voting games: Computational aspects and paradoxes. In *AAAI'10*, pages 805–810, 2010.