# How Altruism Changes Mechanism Design

*Author:*
Sjir Hoeijmakers (BSc Utrecht University) (ANR: 492894)

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science in Operations Research and Management Science

Tilburg School of Economics and Management
Tilburg University

*Supervisors:*
Prof. Dr. Guido Schäfer (CWI and VU University Amsterdam)
Prof. Dr. Peter Borm (Tilburg University)

July 8, 2014

**Abstract**

This work investigates the role altruistic dispositions of players may have in the mechanism design setting. We examine existing and introduce new models in which players range from purely egoistic to fully altruistic in the sense that they care as much about others as about themselves. We interpret and compare these models and their effects, and provide a characterization of truthful mechanisms that are adapted to the altruistic setting. We develop new truthful mechanisms for the models that we have introduced, and prove their favorable properties. We then apply these models and mechanisms to different specific settings in mechanism design and study their effects. In particular, for our most-preferred, VCG mechanisms turn out to remain truthful for any level of altruism among the players. Moreover, a new VCG mechanism satisfying the no positive transfers and individual rationality properties allows us to lift the impossibility result of funding a public project, which was present in standard mechanism design. Lastly, we look into the problem of redistributing payments in the single-item allocation setting, and find that also there altruism may improve results. We conclude that the presence of altruism can lead to significant model-dependent benefits in solving problems in mechanism design, and provide recommendations for future research in this promising area.

## Acknowledgements

First of all I want to thank my main supervisor, Guido Schäfer. He gave me much freedom in pursuing my interests and finding a suitable topic, but was always available to discuss new ideas, provide valuable feedback and give direction whenever necessary. Moreover, it were his inspiring lectures on discrete optimization and his explanation of algorithmic game theory that led me to mechanism design and altruism in the first place.
Secondly, I want to thank my supervisor in Tilburg, Peter Borm, for his willingness to take on this position, and for making time to listen and comment on an exposition of my ideas from a new perspective. This allowed me to get a better overview of and new perspectives on those ideas myself.
I next want to thank CWI for providing me with an internship opportunity in a stimulating and enjoyable environment. I want to thank especially also the people at CWI, and in particular my 'roommates' Mona Rahn and Sandra Hommersom, for making my time there a very memorable and fun experience.
Lastly, I want to thank my parents, without whom I obviously would not have been able to write this thesis, and who have always stimulated and supported me, not only in writing this thesis but in all my efforts.

# Contents

# 1 Introduction

At the very heart of nearly all models in mathematical economics lie the assumptions that human beings are both rational and egoistic. These are certainly very useful assumptions: they hugely simplify analysis and allow economists to make strong predictions on the outcome of economic situations. However, especially when they concern behavior of individuals, often these predictions reflect more what the outcomes 'should' be according to these assumptions than actual reality. In other words: the assumptions are often not justified. This has become all the clearer in the past decades through the advent of the field of behavioral economics. Empirical research there has shown that not only the rationality assumption (Kahneman, 2011) but also the egoism assumption (Andreoni & Miller, 2002; Charnes & Rabin, 2002) often 'fails'.

Mechanism design is in particular a branch that relies heavily on assumptions of egoism. In fact, it is its central goal to counter the negative effects of egoism in a group context. When a group of people has to make a choice, or someone has to make a choice for a group, simply asking the (egoistic) people how much they like each alternative might motivate them to over- or understate their preferences for certain alternatives. In mechanism design, *mechanisms* make sure that extra incentives are provided to the 'egoists' involved, so that they will tell the truth and the best decision can be made. In some sense, the egoism of the people is used to make them act in the interest of the group after all. Of course, this comes at a cost: often payments need to be made by or to the people involved. Mechanism design is the study of finding the 'best' mechanisms that incur the least 'damage' to the group or the person making the decision according to certain criteria.

To be sure, if people would be fully altruistic in that their preferences are aligned with the group's interests, there would be no need for mechanism design. But what happens if they reside in the large spectrum between 'full' altruism and egoism, that is, when they care about others but not as much as about themselves? It seems a reasonable guess that nearly all human beings fall into this category, so motivation for research in this direction is clear.

This thesis contributes to this research by studying different models of altruism, developing new mechanisms that are adapted to altruism, and examining the impact these mechanisms can have in solving problems that are studied in mechanism design. We conclude that, though dependent on the model we use, altruism often has a rather large beneficial influence if 'exploited' in the right way (i.e. by an appropriate mechanism).

In the next chapter we shortly expose previous research in this area, which turns out to still be relatively unexplored. In Chapter 3 we provide the necessary preliminaries from mechanism design. In Chapter 4 we introduce the framework in which we model altruism, and provide the specific models that we use.

We furthermore zoom in on issues of interpretation of both the models already present in the literature and the new ones we propose. In Chapter 5 we consider the design of favorable mechanisms for our models. We first provide a general characterization and proof template for our framework, and subsequently apply the latter to the specific models in question. In Chapter 6 we finally apply the models and mechanisms we introduced to two specific problems in a mechanism design setting, and examine the question of whether unknown altruism may also 'harm' or cause in the sense that VCG mechanisms are no longer truthful when it is present.

## 2 Related literature

Although the role of altruism in (algorithmic) game theory has sparked some interest in recent years (see for instance Caragiannis et al. (2010), De Marco & Morgan (2011) and Apt & Schäfer (2014)), literature on incorporating altruism (or its counterpart spite) in mechanism design has been relatively scarce up to this point. Here, we present the work that has been done in this area.

Brandt and Weiss (2002), first of all, show that when spiteful bidders are present, the second-price or Vickrey auction fails in that it loses its favorable property of truthfulness. They do not present a truthful alternative themselves, but provide a motivation to research implications of spite (and altruism) in mechanism design, and to search for such a truthful alternative.

Tang and Sandholm (2012) also consider the single-item auction format and spiteful bidders, and direct themselves to finding revenue-maximizing mechanisms. Their results furthermore extend to the altruistic domain: they model spite and altruism using the *player-oriented model* that we introduce in Chapter 4. In their proposed revenue-optimal mechanism, a player's own valuation for the object to be auctioned may directly influence her payment. Also, bidders may have to pay even if the auctioneer keeps the item, whereas on the other hand losers are sometimes subsidized by the mechanism.

Kucuksenel (2012) also models altruism according to the player-oriented model but studies its implications in the Bayesian setting. He characterizes a class of mechanisms that are *interim efficient*: they lead to outcomes which cannot be unanimously improved upon by the players utility-wise. He furthermore provides two examples of problems in which altruism has a positive effect on overall performance when interim-efficient mechanisms are applied.

Cavallo (2012), lastly, proposes a regret-based model of altruism. In this model, players are $\alpha$-altruistic if they are willing to sacrifice up to $\alpha$ of their potential utility if that improves the aggregate egoistic utility. He also treats a 'proportional' variant, in which $\alpha$ is not a fixed value but a percentage of the potential utility of a player. In his paper, he uses redistribution to come up with strongly budget-balanced mechanisms (i.e. no net payments are made to or by the mechanism) for the single-item allocation setting when players are at least 'mildly' altruistic. We come back to his model in Chapter 4, and study the implication of our models to the latter setting in Chapter 6.

# 3 Mechanism design

In this chapter, we provide the formal groundwork of mechanism design upon which our later chapters are built. In the first section we introduce the necessary notation. In the second section we define generally favorable properties for mechanisms to satisfy. And in the third section, lastly, we present a class of mechanisms that constitutes the probably most important (or at least most positive) result in mechanism design: VCG mechanisms.
From here on, we will talk about a (fictive) *designer* that develops the mechanism to achieve some specified goal, and that receives and distributes the payments that the mechanism dictates to be made. The designer may in some situations represent the group of players themselves, and in others an entity (e.g. a government) making a choice on behalf of this group, or any other person that wants to take the preferences of the group into account when making a choice.

## 3.1 Framework and notation

We have a group of $n$ players that are labeled by integers from 1 to $n$ in player set $N$. There is a set of alternatives $A$, and one alternative $a \in A$ needs to be chosen. Each player $i \in N$ has a *valuation function* $v_i : A \rightarrow \Re$ that is only known to herself. This valuation function specifies for each alternative the utility this player derives from it, independently from the other players' preferences. For each player $i$ we define $V_i$ to be the set of possible valuation functions (*valuation space*) for this player, which is known by the designer. For a specific alternative $a \in A$, we will call $v_i(a)$ player $i$'s *valuation* of this alternative. We will refer to any utility that arises as a consequence of valuations of alternatives as *value* that is created, while payments are said to transfer this value to other individuals (players or the designer).

We use the following notation for increased clarity. For any function $x$ we write $x$ instead of $x(\cdot)$ when referring to the function itself instead of any entity in its image. For any $n$-dimensional vector of functions $\boldsymbol{x}$ and any $i \in N$ we define $\boldsymbol{x}_{-i}$ such that $\boldsymbol{x} = (x_i, \boldsymbol{x}_{-i})$. With respect to the valuation spaces we will also write $\mathbf{V}$ instead of $V_1 \times \cdots \times V_n$ and $\mathbf{V}_{-i}$ instead of $V_1 \times \ldots V_{i-1} \times V_{i+1} \cdots \times V_n$.

We are now ready to define the notion of a mechanism formally:

**Definition 3.1 (Mechanisms)** A (direct revelation) *mechanism* $(f, \boldsymbol{p})$ is specified by a *social choice function* $f : \mathbf{V} \rightarrow A$ and an $n$-dimensional vector of *payment functions* $\boldsymbol{p}$ such that $p_i : \mathbf{V} \rightarrow \Re$ for all $i \in N$.

This should be interpreted as follows. The mechanism collects a reported valuation function $\bar{v}_i$ from each player $i \in N$. Based on the reports, it chooses an alternative $f(\bar{\boldsymbol{v}})$, where $\bar{\boldsymbol{v}}$ is the $n$-dimensional vector of reports. It further decides on payments $p_i(\bar{\boldsymbol{v}})$ to be made by each player $i$ to the mechanism.

We assume what are called *quasilinear preferences* for the players, meaning that for each player $i$ her valuation $v_i(a)$ of any alternative $a \in A$ can be expressed in the currency of the payments. The total utility of a player is then the difference between the player's valuation of the alternative that is chosen and the payment made by the player to the mechanism. This allows us to define the final utility of a player $i$, dependent on the reported valuations by all players. We will call this the *standard utility* of a player to distinguish it from the other types of utility that we will define in the next chapter to allow for altruistic feelings. For now, players are assumed to be egoistic in the sense that they only care about maximizing the utility function below:

**Definition 3.2 (Standard utility)** Given a mechanism $(f, \boldsymbol{p})$ and private valuation $v_i$, the function $u_i : \mathbf{V} \to \Re$ such that

$$u_i(\bar{\boldsymbol{v}}) = v_i(f(\bar{\boldsymbol{v}})) - p_i(\bar{\boldsymbol{v}})$$

specifies the *standard utility* experienced by player $i$ as a function of the reported valuations of all players.

Whereas maximizing $u_i$ is the goal of an individual player $i$, the social choice function $f$ reflects the goal the designer of the mechanism wants to achieve. It optimizes a certain function of the reported valuation functions of the players over the available alternatives. We will call the function to be maximized by a certain mechanism the *design objective*:

**Definition 3.3 (Design objective)** The *design objective* $D^{\bar{\boldsymbol{v}}} : A \to \Re$ of a mechanism $(f, \boldsymbol{p})$ is the function that $f$ maximizes over the alternatives:

$$f(\bar{\boldsymbol{v}}) \in \arg\max_{a \in A} D^{\bar{\boldsymbol{v}}}(a).$$

Probably the most natural choice for a design objective is the *social welfare*:

**Definition 3.4 (Social welfare)** Given the vector of reported valuation functions $\bar{\boldsymbol{v}}$ of the players, the *social welfare* $SW : A \to \Re$ is specified by

$$SW(a) = \sum_{i \in N} \bar{v}_i(a).$$

For the mechanism to meet its purpose, the designer needs to find a way to ensure that the reported valuations coincide with the real valuation functions of the players. In other words, he needs to provide incentives to the players to tell the truth. This is what the payment functions $\boldsymbol{p}$ are intended for. In the next section we shall introduce formally the property of a mechanism that it makes that players derive maximum utility from truth-telling, and therefore have no reason to lie.

## 3.2 Favorable properties of mechanisms

As already alluded to in the previous section, the most important property a mechanism needs to satisfy for it to be useful is that it makes players want to tell the truth. A mechanism that does so is called *truthful*:

**Definition 3.5 (Truthfulness)** A mechanism $(f, \boldsymbol{p})$ is called truthful if for every player $i \in N$, for any vector of reported valuations $\bar{\boldsymbol{v}} \in \mathbf{V}$ and given the true valuation $v_i \in V_i$, we have that

$$u_i(v_i, \bar{\boldsymbol{v}}_{-i}) \geq u_i(\bar{\boldsymbol{v}}).$$

In other words, under a truthful mechanism it is always better for a player to report her true valuation $v_i$ than any other valuation $\bar{v}_i$, no matter what the other players report or what their true valuations are. This is a very strong property, but without it there might be reasons for people to lie, and there would be no guarantee that the truly preferred alternative (for the designer) is chosen by the mechanism.

A second favorable property of a mechanism is that the designer never has to pay any player any money:

**Definition 3.6 (No positive transfers)** A mechanism $(f, \boldsymbol{p})$ makes *no positive transfers* (NPT) if for every player $i \in N$ and all $\bar{\boldsymbol{v}} \in \mathbf{V}$ we have that

$$p_i(\bar{\boldsymbol{v}}) \geq 0.$$

This makes sense especially when valuations are nonnegative. The same holds for the last property we treat here, which ensures that no player has reason not to participate in the mechanism (if she has a choice to opt out):

**Definition 3.7 (Individual rationality)** A mechanism $(f, \boldsymbol{p})$ satisfies individual rationality (IR) if for every player $i \in N$, for all reported valuations $\bar{\boldsymbol{v}}_{-i} \in \mathbf{V}$ and given the true valuation $v_i \in V_i$, we have that

$$u_i(v_i, \bar{\boldsymbol{v}}_{-i}) \geq 0.$$

## 3.3 VCG mechanisms

One of the most celebrated results in mechanism design is the finding of a set of truthful mechanisms that maximizes the social welfare $SW$. This set of mechanisms was named after William Vickrey, Edward Clarke and Theodore Groves, who all had a role in its discovery (Clarke, 1971; Groves, 1973; Vickrey, 1961):

**Definition 3.8 (VCG mechanisms)** A mechanism $(f, \boldsymbol{p})$ is called a *Vickrey-Clarke-Groves* (VCG) *mechanism* if the following two conditions are satisfied:

- $f(\bar{\boldsymbol{v}}) \in \arg\max_{a \in A} \sum_{i \in N} \bar{v}_i(a)$
- For all $i \in N$, for some function $h_i : \mathbf{V}_{-i} \to \Re$ we have that

$$p_i(\bar{\boldsymbol{v}}) = h_i(\bar{\boldsymbol{v}}_{-i}) - \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})).$$

**Theorem 3.1** All VCG mechanisms are truthful. (Clarke, 1971; Groves, 1973; Vickrey, 1961)

There is a special choice of the functions $h_i$ that furthermore guarantees both NPT and IR to be satisfied whenever valuations are nonnegative. This payment rule is named after its inventor, Edward Clarke (Clarke, 1971):

**Definition 3.9 (Clarke pivot rule)** We say that a VCG mechanism $(f, \boldsymbol{p})$ uses the *Clarke pivot rule* if for every player $i \in N$ we have that

$$h_i(\bar{\boldsymbol{v}}_{-i}) = \sum_{j \in N \setminus \{i\}} v_j(a^{-i})$$

for some $a^{-i} \in \arg\max_{a \in A} \sum_{j \in N \setminus \{i\}} v_j(a)$.

$a^{-i}$ can here be interpreted as an alternative that would maximize social welfare if player $i$ would not be present. With the VCG payments and this choice of $h_i$, player $i$ hence pays exactly for how much his presence negatively affects the utility that other players derive from the alternative that is chosen.

**Lemma 3.1** The VCG mechanism that uses the Clarke pivot rule satisfies NPT. Furthermore, whenever for all $i \in N$, $v_i \in V_i$ and $a \in A$ we have that $v_i(a) \geq 0$, it also satisfies IR. (Clarke, 1971)

With these positive results in mind we now proceed to modeling altruism. We will see how its presence can make for even more well-behaving mechanisms.

# 4 Models of altruism

There are many ways to model altruism in mechanism design, and the choice among these is non-trivial. Different models may result from different beliefs held about altruism or may correspond to different situations in real life, and may yield very different properties for mechanisms. In this chapter we define the models that will be studied in the rest of this work. More specifically, in the first section we define a very general way of incorporating nonstandard preferences into the utility functions of players. In the second section, we introduce two new ways to model altruism that fit into this general framework. In the third section, we interpret the introduced models and compare them with the models in the existing literature.

## 4.1 Utility framework

Let us start by introducing the framework in which we will model altruistic feelings and actions. Our method is to adapt the utility function of the players by adding an extra term representing the altruistic feelings perceived by a player. This is motivated by the fact that in reality altruistic actions often go together with a truly experienced positive feeling, which one may interpret as some form of utility. With this approach, we can furthermore keep the assumption that all players will seek to maximize their utility function individually. We will call he general framework that captures this idea the *utility framework*:

**Definition 4.1 (Utility framework)** Given a mechanism $(f, \boldsymbol{p})$, a private valuation function $v_i$ and an *altruistic disposition function* $g_i : \Re^{n-1} \times \Re^n \to \Re$, the *utility framework* specifies the utility $u_i^{g_i} : \mathbf{V} \to \Re$ of a player $i \in N$ to be:

$$u_i^{g_i}(\bar{\boldsymbol{v}}) = v_i(f(\bar{\boldsymbol{v}})) - p_i(\bar{\boldsymbol{v}}) \ + g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})).$$

The altruistic disposition function $g_i$ of a player $i$ is assumed to be known by the designer. It does not depend on the players' true valuations of the chosen alternative, but only on the reported ones. This reflects the intuition that the positive feelings that go with altruism originate from *beliefs* about the experiences of others rather than from their true experiences. $g_i$ is further allowed to depend on $p_i$ because the payments of player $i$ are not necessarily fully lost to her: if she has altruistic feelings towards the designer or anyone who receives the payments in the end, she will apart from losing utility also gain utility from paying.

Here we use the function $g_i$ to represent the altruistic feelings of player $i$. Note however that in its generality it can represent any utility that player $i$ perceives as a result of the valuations of and payments by other players. Our results for this model hence extend to other forms of 'other-regarding' utilities, such as spite and jealousy.

We should lastly note that specifying and adding the above term $g_i$ to the utility functions of players, as we will do here, is not the only possible way to model altruism. Alternatively, we could restrict the utility function to certain shapes by making assumptions on the behavior it produces. This is the approach Cavallo (2012) adopts in his regret-based model of altruism. We will revisit the interpretation of the different ways to model altruism in Section 4.3.

### 4.2 Proposed models

Let us define the two models that we propose, the *welfare-oriented model* and the *omnistic model*. We also state the *player-oriented model* that is sometimes adopted in the literature (Tang & Sandholm, 2012; Kucuksenel, 2012) and the *regret-based model*, which was introduced by Cavallo (2012):

**Definition 4.2 (Welfare-oriented model)** Given a mechanism $(f, \boldsymbol{p})$, a private valuation function $v_i$ and an *altruism level* $\alpha_i \in [0,1]$, in the *welfare-oriented model*

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

so that the utility $u_i^w : \mathbf{V} \to \Re$ of a player $i \in N$ is given by:

$$u_i^w(\bar{\boldsymbol{v}}) = v_i(f(\bar{\boldsymbol{v}})) - p_i(\bar{\boldsymbol{v}}) \ + \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})).$$

**Definition 4.3 (Omnistic model)** Given a mechanism $(f, \boldsymbol{p})$, a private valuation function $v_i$ and an *altruism level* $\alpha_i \in [0,1]$, in the *omnistic model*

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) + \alpha_i p_i(\bar{\boldsymbol{v}})$$

so that the utility $u_i^o : \mathbf{V} \to \Re$ of a player $i \in N$ is given by:

$$u_i^o(\bar{\boldsymbol{v}}) = v_i(f(\bar{\boldsymbol{v}})) - (1-\alpha_i) p_i(\bar{\boldsymbol{v}}) \ + \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})).$$

**Definition 4.4 (Player-oriented model)** Given a mechanism $(f, \boldsymbol{p})$, a private valuation function $v_i$ and an *altruism level* $\alpha_i \in [0,1]$, in the *player-oriented model*

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = \alpha_i \sum_{j \in N \setminus \{i\}} (\bar{v}_j(f(\bar{\boldsymbol{v}})) - p_j(\bar{\boldsymbol{v}}))$$

so that the utility $u_i^p : \mathbf{V} \to \Re$ of a player $i \in N$ is given by:

$$u_i^p(\bar{\boldsymbol{v}}) = v_i(f(\bar{\boldsymbol{v}})) - p_i(\bar{\boldsymbol{v}}) \ + \alpha_i \sum_{j \in N \setminus \{i\}} (\bar{v}_j(f(\bar{\boldsymbol{v}})) - p_j(\bar{\boldsymbol{v}})).$$

The $p$ used to label the utility of the player-oriented model should not be confused with $p_i$ or $\boldsymbol{p}$, which we use to refer to payments. In all models, when $\alpha_i = \lambda$ for some $\lambda \in [0,1]$, we say that player $i$ is $\lambda$-altruistic.

## 4.3 Interpretation

In this section, we compare the three models in the utility framework above with each other and with the regret-based model proposed by Cavallo (2012) on the basis of their interpretation. Let us start by informally explaining the latter, as it does not fit the utility framework.

In the regret-based model (Cavallo, 2012), a player is $\lambda$-altruistic if she weakly prefers reporting her true valuation function to making any other report whenever two conditions are satisfied. Firstly, reporting the truth should mean giving up at maximum $\lambda$ of the standard utility she could obtain by making this other report. Secondly, the truthful report should result in an aggregate standard utility of all the players weakly larger than the aggregate standard utility obtained by making the other report. Recall that the aggregate standard utility is different from the social welfare, in that the former includes the payments that are made. In this sense, Cavallo's regret-based model has some similarity to the player-oriented model: players do not have any altruistic feelings towards the designer, or whoever receives the payments.
The model deserves the name 'regret-based' from the fact that players are assumed to experience altruism dependent on the standard utility they *could have* obtained. It does not fit the utility framework outlined above, as $g_i$ is not allowed to depend on this contextual fact. Note that the regret-based model does not specify a term to be added to the utility functions, but restricts the allowed utility functions by means of referring to the behavior that should result from them.
The regret-based model is intuitive in the sense that it sees altruism as the willingness to sacrifice some of ones own standard utility for that of others. The most important objection one could make, however, is that it takes the maximum sacrifice to be a fixed number, in that it is independent of *by how much* the sacrifice improves the standard utility of others. This is not the case in the other three models, in which players experience non-standard utility proportional to how much (standard) utility others experience.

The player-oriented, welfare-oriented and omnistic model all add to the utility of player $i$ a term that depends linearly on the other players' valuations of the chosen alternative. For all three, we will study mechanisms when values of $\alpha_i$ range from 0 to 1. $\alpha_i = 0$ corresponds to the 'full egoism' that is present in the standard utility model, and $\alpha_i = 1$ represents 'full altruism' in the sense that a player cares as much about the valuations of others as about those of herself. Note though that many results that will follow may be easily extended to the case of spiteful players ($\alpha_i < 0$) or players that care even more about others than about themselves ($\alpha_i > 1$).

The three models in the utility framework mainly differ in how they handle the payments that are made. In the player-oriented model, firstly, the players do not perceive any feelings of altruism towards the designer or to what she represents. The payments to the designer are fully lost to the players, and any transfer from the designer to the players is experienced as a true 'creation' of utility, non-distinct from when valuations would be higher by the same amount. Consequently, players are indifferent between having valuations of 0 while not having to pay anything and having high valuations while they have to pay an amount equal to these valuations. The second situation is however better utility-wise for the 'universe' as a whole, because the payments are not lost: the utility that was created by the players is simply redistributed to other people. The player-oriented model hence assumes a kind of *directed* altruism, as opposed to the *non-directed* or truly *uniform* altruism that we are mostly studying in this work. Moreover, the utility model should not be used when the designer represents the group of players themselves (e.g. if it is a government), because in this case the payments in some way or another end up with the players again. The same criticism holds for the regret-based model, which as noted before cares about payments in a similar way.

In the welfare-oriented model, secondly, players do not care about payments made by or to other players, but care fully about the payments they make themselves. Altruism here corresponds to a willingness to contribute to the creation of value in the form of valuations of alternatives. It is certainly non-directed, as a player here does not care at all where the extra value is created. The welfare-oriented model is intuitive in most situations, but care should be taken in applying it to situations in which the designer functions as a 'mediator' (when she both is paid by players and pays players), as the following example shows.

**Example 4.1** Consider two fully altruistic ($\alpha_1 = \alpha_2 = 1$) truth-telling players in the welfare-oriented model, and say that a mechanism has chosen an alternative $a$ and payment vector $\boldsymbol{p}$ so that $v_1(a) = 0$, $v_2(a) = 2$, $p_1 = -1$ and $p_2 = 1$. When the players would have been egoistic, they would have both derived a utility of 1 from this situation, but as they are altruistic, we should probably expect them to share in the other player's joy and both derive a bit more. In any case, there seems to be no reason for one of the two player's to perceive a larger amount of utility than other, since both derive an equal standard utility from the situation. The predictions of the welfare-oriented model however contradict this logic, stating that $u_1 = 3$ while $u_2 = 1$. In some sense Player 1, who is receiving the money, is able to count the value that Player 2 creates through his valuations twice: once through her 'altruistic' feelings and once through the payment she receives. Player 2, however, only derives utility from the value that Player 1 creates, and not for the value that she creates herself but has to give to Player 1.

When the designer is not a mediator but is either only paid (i.e. it satisfies the no positive transfers property) or only pays, the welfare-oriented model remains an intuitive and plausible model.

Thirdly and lastly, consider the omnistic model that we propose. In it, a player derives utility from any standard utility that others perceive. Here 'others' includes every possible affected person, so also individuals that are not partaking in the mechanism and may be represented by the designer, who receives the payments that the players make. This is the reason the model is called the 'omnistic' model (omnes = all/everybody). It models a truly non-directed form of altruism, in that a player derives utility to the same extent from any value that is created, no matter with whom it ends up apart from herself. The model is applicable to any situation in which the latter holds. Note that it can easily be adjusted to a more directed type of model by using parameters $\alpha_{ij}$ for the disposition of $i$ towards $j$, and by adding and extra parameter $\alpha_i D$ to represent the disposition towards the designer, which results in the payment discount. The player-oriented model is then a special case of the omnistic model with $\alpha_i D = 0$.

# 5 Altruism-adapted mechanisms

In this chapter, we look into mechanisms that take into account the altruistic feelings of players as modeled in the utility framework. In the first two sections, we derive sufficient conditions for mechanisms to be truthful and satisfy no positive transfers (NPT) and individual rationality (IR) in this general framework. In the third and fourth sections, we will use these conditions to derive favorable mechanisms for both the welfare-oriented and the omnistic model.

## 5.1 Truthfulness

In this section we first provide a direct characterization of truthful mechanisms in the utility framework similar to what is given for the standard utility model in (Nisan et al., 2007, Chapter 9). Thereafter we provide another sufficient condition for truthfulness that will help us in deriving truthful mechanisms and functions as a proof template for their truthfulness. Before we start, note that though we defined the concept of truthfulness in Section 3.2 with respect to standard utility, it extends in a natural way to all utility functions in the utility framework by replacing $u_i$ with $u_i^{g_i}$ in the definition. From here on, then, we will speak of mechanisms being truthful within a certain model.

**Proposition 5.1** A mechanism $(f, \boldsymbol{p})$ is truthful in a model in the utility framework if and only if it satisfies the following two conditions for every player $i$ and all reports of other players $\bar{\boldsymbol{v}}_{-i}$:

- The difference between the altruistic disposition $g_i$ and the payment $p_i$ does not depend directly on the reported $\bar{v}_i$ but only does so through the alternative chosen $f(\bar{\boldsymbol{v}})$. That is, for some function $\mu_i : A \times V_{-i} \to \Re$ we have that
$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) - p_i(\bar{\boldsymbol{v}}) = \mu_i(f(\bar{\boldsymbol{v}}), \bar{\boldsymbol{v}}_{-i})$$

- The mechanism maximizes the utility $u_i^{g_i}$ for player $i$ when this player reports her valuation function truthfully. That is, for any $v_i \in V_i$ we have that
$$f(v_i, \bar{\boldsymbol{v}}_{-i}) \in \arg\max_{a \in A^{\bar{v}_{-i}}} (v_i(a) + \mu_i(a, \bar{\boldsymbol{v}}_{-i}))$$
where $A^{v_{-i}}$ constitutes the range of $f(\cdot, \bar{\boldsymbol{v}}_{-i})$.

*Proof* We start with the if part. Given $\bar{\boldsymbol{v}}_{-i}$, player $i$'s true valuation $v_i$ and any other report $\bar{v}_i'$, denote $a^* = f(v_i, \bar{\boldsymbol{v}}_{-i})$, $a' = f(\bar{v}_i', \bar{\boldsymbol{v}}_{-i})$, $\mu_i^{a*} = \mu(f(v_i, \bar{\boldsymbol{v}}_{-i}), \bar{\boldsymbol{v}}_{-i}))$ and $\mu_i^{a'} = \mu(f(\bar{v}_i', \bar{\boldsymbol{v}}_{-i}), \bar{\boldsymbol{v}}_{-i}))$. Player $i$'s utility when telling the truth is then $v_i(a^*) + \mu_i^{a*}$, which cannot be less than her utility $v_i(a') + \mu_i^{a'}$ when reporting $\bar{v}_i'$, since the mechanism maximizes $i$'s utility, that is,

$$a^* \in \arg\max_{a \in A^{\bar{v}_{-i}}} (v_i(a) + \mu_i(a, \bar{\boldsymbol{v}}_{-i}))$$

Now we prove the only-if part of the first condition. Whenever for some $v_i$ and $\bar{v}_i'$ we have that $f(v_i, \bar{\boldsymbol{v}}_{-i}) = f(\bar{v}_i', \bar{\boldsymbol{v}}_{-i})$ but $\mu(f(\bar{v}_i', \bar{\boldsymbol{v}}_{-i}), \bar{\boldsymbol{v}}_{-i})) > \mu(f(v_i, \bar{\boldsymbol{v}}_{-i}), \bar{\boldsymbol{v}}_{-i}))$,

then a player that has true valuation function $\bar{v}_i$ can increase her utility by misreporting.
Lastly, we prove the only-if part of the second condition. Assume given the true valuation $v_i$ that

$$f(v_i, \bar{\boldsymbol{v}}_{-i}) \notin \arg\max_{a \in A^{\bar{\boldsymbol{v}}_{-i}}} (v_i(a) + \mu_i(a, \bar{\boldsymbol{v}}_{-i}))$$

and take an $a' \in A^{\bar{\boldsymbol{v}}_{-i}}$ such that

$$a' \in \arg\max_{a \in A^{\bar{\boldsymbol{v}}_{-i}}} (v_i(a) + \mu_i(a, \bar{\boldsymbol{v}}_{-i}))$$

By definition of $A^{\bar{\boldsymbol{v}}_{-i}}$ we now have that $a' = f(\bar{v}'_i, \bar{\boldsymbol{v}}_{-i})$ for some $\bar{v}'_i \in V_i$, so that player $i$ can improve on her utility by misreporting. □

The first condition tells us something important about the interaction between a model in the utility framework and the mechanisms that are truthful in it. For note what happens if $g_i$ depends on $p_{-i}$ for each $i$, as it does in the player-oriented model. In such a case, if we want $p_i$ for each $i$ to depend on the *function* $\bar{\boldsymbol{v}}_{-i}$ (that is, not only on $v_{-i}(f(\bar{\boldsymbol{v}}))$), we need it to depend on the function $\bar{v}_i$ as well. If it does not, the first condition above cannot hold, because then on the left-hand side (through $g_i$) a dependence on $\bar{v}_i$ arises that does not exist on the right-hand side. This may severely limit our choice of truthful payment functions, as an addition of an arbitrary function $h_i(\bar{\boldsymbol{v}}_{-i})$ to $p_i$ may now have an influence on truthfulness. Recall that in VCG mechanisms (in the standard utility model), any such additional function $h_i(\bar{\boldsymbol{v}}_{-i})$ could be chosen, and that this freedom was used to come up with the favorable Clarke pivot rule.
Although Proposition 5.1 is a characterization of truthful mechanisms in the utility framework, it does not directly provide us with a recipe to obtain such a mechanism for a given design objective $D$. The sufficient condition that we next provide does this, and we will be able to use it later as a proof template for the truthful mechanisms that we introduce in the coming sections.

**Proposition 5.2** A mechanism $(f, \boldsymbol{p})$ is truthful in any model specified by $g_i$ for all $i \in N$ in the utility framework, if for all $i$, true valuation functions $v_i$ and reports $\bar{\boldsymbol{v}}$, and for some functions $h_i : \mathbf{V}_{-i} \to \Re$, $\gamma_i : \mathbf{V}_{-i} \to \Re^+$ and $\zeta_i : \mathbf{V} \to \Re$ two conditions are satisfied:

$$p_i(\bar{\boldsymbol{v}}) = h_i(\bar{\boldsymbol{v}}_{-i}) + g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) - \zeta(\bar{\boldsymbol{v}}) \tag{1}$$

$$\zeta(\bar{\boldsymbol{v}}) = \gamma_i(\bar{\boldsymbol{v}}_{-i}) D^{(v_i, \bar{\boldsymbol{v}}_{-i})}(f(\bar{\boldsymbol{v}})) - v_i(f(\bar{\boldsymbol{v}})) \tag{2}$$

*Proof* We need to show that $u_i^{g_i}(v_i, \bar{\boldsymbol{v}}_{-i}) \geq u_i^{g_i}(\bar{\boldsymbol{v}})$ for any player $i$, true valuation function $v_i$ and report vector $\bar{\boldsymbol{v}}$. Plugging (1) and (2) into Definition 4.1 gives for the utilities:

$$u_i^{g_i}(v_i, \bar{\boldsymbol{v}}_{-i}) = \gamma_i(\bar{\boldsymbol{v}}_{-i}) D^{(v_i, \bar{\boldsymbol{v}}_{-i})}(f(v_i, \bar{\boldsymbol{v}}_{-i})) + h_i(\bar{\boldsymbol{v}}_{-i})$$

and

$$u_i^{g_i}(\bar{\boldsymbol{v}}) = \gamma_i(\bar{\boldsymbol{v}}_{-i}) D^{(v_i, \bar{\boldsymbol{v}}_{-i})}(f(\bar{\boldsymbol{v}})) + h_i(\bar{\boldsymbol{v}}_{-i})$$

so that

$$
\begin{aligned}
u_i^{g_i}(v_i, \bar{\boldsymbol{v}}_{-i}) - u_i^{g_i}(\bar{\boldsymbol{v}}) &= \gamma_i(\bar{\boldsymbol{v}}_{-i})(D^{(v_i,\bar{\boldsymbol{v}}_{-i})}(f(v_i, \bar{\boldsymbol{v}}_{-i})) - D^{(\bar{v}_i,\bar{\boldsymbol{v}}_{-i})}(f(\bar{\boldsymbol{v}}))) \\
&\geq 0
\end{aligned}
$$

where the final inequality follows from the definition of $D^{\bar{\boldsymbol{v}}}$ and from $\gamma_i(\bar{\boldsymbol{v}}_{-i}) \geq 0$. □

Condition (2) is important to guarantee that the payments can actually be calculated and issued. Without it, they may depend on the true valuation functions $v_i$, which are by definition unknown to the designer. This condition fixes $\gamma_i$ in such a way that the appearance of $v_i$ in $D^{(v_i,\bar{\boldsymbol{v}}_{-i})}$ is cancelled by its appearance in the utility function $u_i^{g_i}$. Note that in mechanism design using the standard utility function and focused on maximizing the social welfare, this happens for $\gamma_i$ being the constant function 1. $h_i - \zeta_i$ then resembles the VCG payments.
On the interpretation side, note especially the positive dependence of $p_i$ on $g_i$. This indicates that the more positive altruistic feelings a player experiences, the less we need to pay this player to report her valuation function truthfully.

With the above conditions we can derive new truthful mechanisms and prove the truthfulness of a given mechanism, which we will do in Sections 5.3 and 5.4 for the welfare-oriented model and the omnistic model respectively. The freedom of choice for the function $h_i$ furthermore allows us to tune our truthful mechanisms to specific needs, and to this tuning we now turn.

## 5.2 NPT and IR

Just as the Clarke pivot rule allows us to do in standard mechanism design, we would like to be able to ensure that a mechanism satisfies NPT and IR when valuation functions are nonnegative. The proof template introduced in the previous section allows us to specify precisely for which choices of $h_i$ a mechanism indeed does this, which we make formal in the following two propositions:

**Proposition 5.3** A truthful mechanism $(f, \boldsymbol{p})$, with payments as specified in Proposition 5.2, satisfies NPT in the utility framework if and only if the following holds for all $i$, true valuations $v_i$ and reported valuations $\bar{\boldsymbol{v}}$:

$$h_i(\bar{\boldsymbol{v}}_{-i}) \geq \zeta(\bar{\boldsymbol{v}}) - g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) \tag{3}$$

*Proof* This follows directly from requiring $p_i(\bar{\boldsymbol{v}}) \geq 0$, substituting condition (1) from Proposition 5.2 and rearranging the result. □

**Proposition 5.4** A truthful mechanism $(f, \boldsymbol{p})$, with payments as specified in Proposition 5.2, satisfies IR in the utility framework if and only if the following holds for all $i$, true valuations $v_i$ and reported valuations $\bar{\boldsymbol{v}}$:

$$h_i(\bar{\boldsymbol{v}}_{-i}) \leq \zeta(\bar{\boldsymbol{v}}) + v_i(f(\bar{\boldsymbol{v}})) \tag{4}$$

*Proof* This follows directly from requiring $u_i^{g_i}(v_i, \bar{\boldsymbol{v}}_{-i}) \geq 0$, substituting condition (1) from Proposition 5.2 and rearranging the result (note that $\bar{v}_i$ equals $v_i$ here). $\square$

We see from (3) and (4) that in principle there is a leeway of choosing $h_i$ of size $v_i(f(\bar{\boldsymbol{v}})) + g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}}))$. $h_i$ can however only depend on $\bar{\boldsymbol{v}}_{-i}$, and not on $\bar{v}_i$ or $v_i$. The latter shows us why we will need the assumption of nonnegative valuations. Without it, there can be no guarantee that NPT and IR are satisfied at the same time. With it, we are left with a playing room of size $g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}}))$, but only in as far as it depends on $\bar{\boldsymbol{v}}_{-i}$, and its dependence on $\bar{v}_i$ (through $\boldsymbol{p}(\bar{\boldsymbol{v}})$) should be nonnegative.
Note that if $g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}}))$ is allowed to be negative itself, we cannot guarantee both NPT and IR. This would happen, for instance, when players are spiteful. When only altruism is present and valuation functions are nonnegative we can expect $g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}}))$ to be nonnegative, since this corresponds to player $i$ deriving utility from other players' (standard) utility.
When $g_i = 0$, as in standard mechanism design, we see that our room to choose $h_i$ is very limited. In fact, it may seem there is no feasible choice for $\delta$ at all, since $\zeta_i$ may also still depend on $\bar{v}_i$ whereas $h_i$ cannot. $\zeta_i$'s dependence on $\bar{v}_i$, however, is only through the alternative that is chosen by the mechanism, as can be seen in (2) from Proposition 5.2. When maximizing social welfare, there turns out to be exactly one choice of $\delta$ that satisfies both inequalities, and this choice is given by the Clarke pivot rule (Definition 3.9). This rule interestingly exploits the appearance of $v_i$ in (4) without explicitly including it in the choice for $h_i$. By choosing the alternative that maximizes social welfare 'without' player $i$ at the left-hand side, (3) is certainly satisfied, for the alternative chosen at the right-hand side maximizes the 'full' social welfare. (4) is satisfied as well, since the addition of $v_i$ to the right-hand side makes this side represent exactly this full social welfare.

## 5.3 Mechanisms for the welfare-oriented model

We are now ready to present favorable mechanisms for our introduced models, starting with the welfare-oriented model in this section. For this model, we provide a class of truthful mechanisms that maximize the social welfare $SW$. We call them *altruism-adjusted VCG mechanisms* for their resemblance to the well-known class:

**Definition 5.1 (AAVCG mechanisms)** A mechanism $(f, \boldsymbol{p})$ is called an *altruism-adjusted Vickrey-Clarke-Groves* (AAVCG) *mechanism* if the following two conditions are satisfied:

- $f(\bar{\boldsymbol{v}}) \in \arg\max_{a \in A} \sum_{i \in N} \bar{v}_i(a)$
- For all $i \in N$, for some function $h_i : \mathbf{V}_{-i} \to \Re$ and given the altruism levels $\alpha_i \in [0, 1]$, we have that

$$p_i(\bar{\boldsymbol{v}}) = h_i(\bar{\boldsymbol{v}}_{-i}) - (1 - \alpha_i) \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})).$$

AAVCG mechanisms reduces to VCG mechanisms when players are fully egoistic. They nicely capture the intuition that we are using them to counter the negative effects of egoistic predispositions. When a player $i$ becomes more altruistic (so that $\alpha_i$ grows), we need to pay her less to have her want to speak the truth about her valuation function. In fact, as we should expect, players require no extra incentive at all when they are fully altruistic ($\alpha_i = 1$).

Let us prove our claim of truthfulness:

**Theorem 5.1** All AAVCG mechanisms are truthful in the welfare-oriented model.

*Proof* Recall from Definition 4.2 that in the welfare-oriented model we have

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

and that our choice of $D^{\bar{\boldsymbol{v}}}$ here yields

$$D^{(v_i, \bar{\boldsymbol{v}}_{-i})}(f(\bar{\boldsymbol{v}})) = v_i(f(\bar{\boldsymbol{v}})) + \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

Plugging these into equations (2) and (1) in Proposition 5.2, taking $\gamma_i = 1$ and substituting $\zeta_i$ gives:

$$p_i(\bar{\boldsymbol{v}}) = h_i(\bar{\boldsymbol{v}}_{-i}) - (1 - \alpha_i) \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

which proves the claim. □

Next we characterize a set of choices for $h_i$ that makes the AAVCG mechanism satisfy both NPT and IR if valuation functions are nonnegative. We call this the *altruism-adjusted Clarke pivot rule* for obvious reasons:

**Definition 5.2 (Altruism-adjusted Clarke pivot rule)** We say that an AAVCG mechanism $(f, \boldsymbol{p})$ uses the *altruism-adjusted Clarke pivot rule* if for every player $i \in N$ there is some $c_i \in [0, \alpha_i]$ such that

$$h_i(\bar{\boldsymbol{v}}_{-i}) = (1 - \alpha_i + c_i) \sum_{j \in N \setminus \{i\}} v_j(a^{-i})$$

for some $a^{-i} \in \arg\max_{a \in A}(\sum_{j \in N \setminus \{i\}} v_j(a))$.

We see that altruism-adjusted Clarke pivot rule allows for a set of mechanisms (parametrized by $c_i$ for all $i$) instead of just one mechanism, and that the size of this set grows with the altruism levels $\alpha_i$ of the players. When no altruism is present, the rule reduces to the Clarke pivot rule.

We formalize the rule's favorable properties in the following lemma:

**Lemma 5.1** Whenever for all $i \in N$, $v_i \in V_i$ and $a \in A$ we have that $v_i(a) \geq 0$, an AAVCG mechanism that uses the altruism-adjusted Clarke pivot rule satisfies NPT and IR in the welfare-oriented model.

*Proof* We will use Proposition 5.2:

$$\begin{aligned}
h_i(\bar{\boldsymbol{v}}_{-i}) &= (1-\alpha_i+c_i)\sum_{j\in N\setminus\{i\}}\bar{v}_j(a^{-i}) \\
&\geq (1-\alpha_i+c_i)\sum_{j\in N\setminus\{i\}} v_j(f(\bar{\boldsymbol{v}})) \\
&= \zeta(\bar{\boldsymbol{v}}) - g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) + c_i \sum_{j\in N\setminus\{i\}} v_j(f(\bar{\boldsymbol{v}})) \\
&\geq \zeta(\bar{\boldsymbol{v}}) - g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}}))
\end{aligned}$$

so that by (3) from Proposition 5.3 NPT is satisfied. Here the first inequality follows from the definition of $a^{-i}$. The second equality follows from substituting condition (2) from Proposition 5.2. with $\gamma = 1$ and by definition of $g_i$ in the welfare-oriented model. The third inequality follows from nonnegativity of $v_j$ for all $j$.
Secondly, using Proposition 5.4:

$$\begin{aligned}
h_i\left(\bar{\boldsymbol{v}}_{-i}\right) &= (1-\alpha_i+c_i)\sum_{j\in N\setminus\{i\}}\bar{v}_j\left(a^{-i}\right) \\
&\leq (1-\alpha_i+c_i)\sum_{j\in N}\bar{v}_j\left(f\left(\bar{\boldsymbol{v}}\right)\right) \\
&\leq \sum_{j\in N}\bar{v}_j\left(f\left(\bar{\boldsymbol{v}}\right)\right) \\
&= \zeta\left(\bar{\boldsymbol{v}}\right) + v_i\left(f\left(\bar{\boldsymbol{v}}\right)\right)
\end{aligned}$$

so that by (4) from Proposition 5.4 IR is satisfied. The first inequality follows from the definition of $f(\bar{\boldsymbol{v}})$ and from $v_i \geq 0$, the second from nonnegativity of $v_j$ for all $j$ and the second equality follows from (2) from Proposition 5.2 with $\gamma = 1$ and from $\bar{v}_i = v_i$, which holds here. $\square$

### 5.4 Mechanisms for the omnistic model

The social welfare is an intuitive design objective, especially when players are assumed to be fully egoistic. When players are altruistic, however, one might argue that we should take into account the altruistic preferences of players alongside their 'egoistic' ones (the valuation functions). In the omnistic model, a design objective in accordance with this line of thought is the *omnistic social welfare* $SW^o$, which is defined as follows:

**Definition 5.3 (Omnistic social welfare)** Given the vector of reported valuation functions $\bar{\boldsymbol{v}}$ of the players, the *omnistic social welfare* $SW^o : A \to \Re$ is specified by

$$SW^o(a) = \sum_{i\in N}\left(1 + \sum_{j\in N\setminus\{i\}}\alpha_j\right)\bar{v}_i(a).$$

The omnistic social welfare is comparable to the social welfare in that both can be seen as the sum of the utilities of the players plus an extra component representing the utility of the designer, who is modeled as a 'typical' player. For the social welfare and the standard utility model, the designer utility (let us call it $u_d$ for the moment) is then the sum of the payments of all the players to him, so that:

$$\begin{aligned} SW &= \sum_{i \in N} u_i + u_d \\ &= \sum_{i \in N} (v_i - p_i) + \sum_{i \in N} p_i \\ &= \sum_{i \in N} v_i. \end{aligned}$$

In the omnistic model a typical player will generally like receiving payments less than in the standard utility model. This is because receiving the payments means that someone else has had to pay and has thereby lost (standard) utility. One might argue that a choice of

$$u_d^o = \left(1 - \frac{\sum_{j \in N} \alpha_j}{n}\right) \sum_{i \in N} p_i$$

is then most intuitive, but this results in a design objective that depends on the payments made, which is problematic if one uses these payments to establish truthfulness. We therefore approximate the designer utility by

$$u_d^o = \sum_{i \in N} (1 - \alpha_i) p_i$$

which gives:

$$\begin{aligned} SW^o &= \sum_{i \in N} u_i^o + u_d^o \\ &= \sum_{i \in N} \left( v_i - (1 - \alpha_i)\, p_i + \alpha_i \sum_{j \in N \setminus \{i\}} v_j \right) + \sum_{i \in N} (1 - \alpha_i)\, p_i \\ &= \sum_{i \in N} \left( 1 + \sum_{j \in N \setminus \{i\}} \alpha_j \right) v_i. \end{aligned}$$

We are now ready to introduce a new type of mechanism, which we call *altruism-balanced*. Altruism-balanced mechanisms are mechanisms that both maximize the omnistic social welfare and are truthful in the omnistic model. They are defined as follows:

**Definition 5.4 (Altruism-balanced mechanisms)** A mechanism $(f, \boldsymbol{p})$ is called an *altruism-balanced mechanism* if the following two conditions are satisfied:

- $f(\bar{\boldsymbol{v}}) \in \arg\max_{a \in A} SW^o(a)$
- For all $i$, for some function $h_i : \mathbf{V}_{-i} \to \Re$ we have that

$$p_i(\bar{\boldsymbol{v}}) = h_i(\bar{\boldsymbol{v}}_{-i}) - (1-\alpha_i)^{-1} \left( \sum_{j \in N \setminus \{i\}} \frac{1 + \sum_{k \in N \setminus \{j\}} \alpha_k}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} - \alpha_i \right) \bar{v}_j(f(\bar{\boldsymbol{v}})).$$

The quotient in the sum above makes for a fundamental new feature with respect to VCG mechanisms. For any other player $j$, player $i$ will receive more than $j$'s valuation in payments if player $i$ is more altruistic than $j$ and less if $i$ is less altruistic than $j$. This stands in interesting contrast with the AAVCG mechanisms, in which the more altruistic you were, the less you would receive. Note also that in the latter, a player's 'earnings' depended on her *absolute* altruism level, whereas in altruism-balanced mechanisms they depend on her altruism level *relative* to the altruism levels of the other players.
Altruism-balanced mechanisms can, just as AAVCG mechanisms, be seen as a generalization of VCG mechanisms to when players can be altruistic. When $\alpha_i = 0$ for all $i$, altruism-balanced mechanisms reduce to VCG mechanisms. Moreover, this happens in any case of universal altruism ($\alpha_i = \alpha$ for all $i$), which is explained by the fact that in such cases maximizing omnistic social welfare is equivalent to maximizing social welfare.

Let us proceed to prove that altruism-balanced mechanisms are indeed truthful:

**Theorem 5.2** All altruism-balanced mechanisms are truthful in the omnistic model.

*Proof* We choose $\gamma_i$ to be the following:

$$\gamma_i = \left( 1 + \sum_{l \in N \setminus \{i\}} \alpha_l \right)^{-1}$$

This makes that:

$$\begin{aligned} \zeta_i(\bar{\boldsymbol{v}}) = & \frac{\left(1 + \sum_{l \in N \setminus \{i\}} \alpha_l\right) v_i(f(\bar{\boldsymbol{v}})) + \sum_{j \in N \setminus \{i\}} \left(1 + \sum_{k \in N \setminus \{j\}} \alpha_k\right) \bar{v}_j(f(\bar{\boldsymbol{v}}))}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} \\ & - v_i(f(\bar{\boldsymbol{v}})) \\ = & \sum_{j \in N \setminus \{i\}} \frac{1 + \sum_{k \in N \setminus \{j\}} \alpha_k}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} \bar{v}_j(f(\bar{\boldsymbol{v}})). \end{aligned}$$

Recall from Definition 4.3 that in the omnistic model we have

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) + \alpha_i p_i(\bar{\boldsymbol{v}})$$

Plugging the above two equations into condition (1) in Proposition 5.2 yields that we need for some $h_i$:

$$\begin{aligned}
p_i(\bar{\boldsymbol{v}}) &= h_i(\bar{\boldsymbol{v}}_{-i}) + \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) + \alpha_i p_i(\bar{\boldsymbol{v}}) \\
&\quad - \sum_{j \in N \setminus \{i\}} \frac{1 + \sum_{k \in N \setminus \{j\}} \alpha_k}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} \bar{v}_j(f(\bar{\boldsymbol{v}})) \\
&= h_i(\bar{\boldsymbol{v}}_{-i}) - \left( \sum_{j \in N \setminus \{i\}} \frac{1 + \sum_{k \in N \setminus \{j\}} \alpha_k}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} - \alpha_i \right) \bar{v}_j(f(\bar{\boldsymbol{v}})) + \alpha_i p_i(\bar{\boldsymbol{v}}).
\end{aligned}$$

Rearranging this and defining $h'_i(\bar{\boldsymbol{v}}_{-i}) = (1-\alpha_i)^{-1} h'_i(\bar{\boldsymbol{v}}_{-i})$ gives us:

$$p_i(\bar{\boldsymbol{v}}) = h'_i(\bar{\boldsymbol{v}}_{-i}) - (1-\alpha_i)^{-1} \left( \sum_{j \in N \setminus \{i\}} \frac{1 + \sum_{k \in N \setminus \{j\}} \alpha_k}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} - \alpha_i \right) \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

which proves the claim. □

Lastly, we will introduce the *altruism-balanced payment rule* to find altruism-balanced mechanisms that satisfy NPT and IR when valuations are nonnegative.

**Definition 5.5 (Altruism-balanced payment rule)** We say that an altruism-balanced mechanism $(f, \boldsymbol{p})$ uses the *altruism-balanced payment rule* if for each player $i$ there is some $c_i \in [0, \alpha_i]$ such that

$$h_i(\bar{\boldsymbol{v}}_{-i}) = (1-\alpha_i)^{-1} \left( \sum_{j \in N \setminus \{i\}} \frac{1 + \sum_{k \in N \setminus \{j\}} \alpha_k}{1 + \sum_{l \in N \setminus \{i\}} \alpha_l} - \alpha_i + c_i \right) \bar{v}_j(f(a^{-i}))$$

and

$$a^{-i} \in \arg\max_{a \in A} \sum_{j \in N \setminus \{i\}} \left( 1 + \sum_{k \in N \setminus \{i,j\}} \alpha_k - \alpha_i \left( \sum_{l \in N \setminus \{i\}} \alpha_l \right) \right) \bar{v}_j(a).$$

We see that, just as in the welfare-oriented model, the set of truthful mechanisms satisfying NPT and IR under nonnegative valuations grows with the altruism levels of the players. The special case $\alpha_i = 0$ for all $i$ again corresponds to the Clarke pivot rule. When $\alpha_i = \alpha$ for all $i$ on the other hand, the altruism-balanced pivot rule specifies a larger set of payment rules that contains the Clarke pivot rule. This was to be expected, as in such a case all altruism-balanced mechanisms are VCG mechanisms, which we saw earlier in this section.

We also see that $a^{-i}$ can again be interpreted as maximizing the design objective 'without' player $i$, though with an extra 'discount' equal to $\alpha_i \left( \sum_{l \in N \setminus \{i\}} \alpha_l \right) \sum_{j \in N \setminus \{i\}} \bar{v}_j(a)$.

Lastly, note that the more altruistic a player is *relative* to the other players, the more she generally pays, and vice versa. This is because the altruism-dependent quotient that we talked about earlier also appears in $h_i$. And since in most cases (though certainly not always) we have that

$$\bar{v}_j(f(a^{-i})) \geq \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

a relatively high $\alpha_i$ will mostly increase payments rather than decrease them.

We will finish this chapter by proving the claim we made with respect to the NPT and IR properties and the altruism-balanced payment rule:

**Lemma 5.2** Whenever for all $i \in N$, $v_i \in V_i$ and $a \in A$ we have that $v_i(a) \geq 0$, an altruism-balanced mechanism that uses the altruism-balanced payment rule satisfies NPT and IR in the omnistic model.

*Proof* We prove NPT directly from Definition 3.6:

$$\begin{aligned}
p_i(\bar{\boldsymbol{v}}) &= h_i(\bar{\boldsymbol{v}}_{-i}) - (1-\alpha_i)^{-1} \sum_{j \in N\setminus\{i\}} \left( \frac{1 + \sum_{k \in N\setminus\{j\}} \alpha_k}{1 + \sum_{l \in N\setminus\{i\}} \alpha_l} - \alpha_i \right) \bar{v}_j(f(\bar{\boldsymbol{v}})) \\
&= (1-\alpha_i)^{-1} \sum_{j \in N\setminus\{i\}} \left( \frac{1 + \sum_{k \in N\setminus\{j\}} \alpha_k}{1 + \sum_{l \in N\setminus\{i\}} \alpha_l} - \alpha_i \right) (\bar{v}_j(f(a^{-i})) - \bar{v}_j(f(\bar{\boldsymbol{v}}))) \\
&\quad + c_i(1-\alpha_i)^{-1} \sum_{j \in N\setminus\{i\}} \bar{v}_j(f(a^{-i})) \\
&\geq (1-\alpha_i)^{-1} \sum_{j \in N\setminus\{i\}} \left( \frac{1 + \sum_{k \in N\setminus\{j\}} \alpha_k}{1 + \sum_{l \in N\setminus\{i\}} \alpha_l} - \alpha_i \right) (\bar{v}_j(f(a^{-i})) - \bar{v}_j(f(\bar{\boldsymbol{v}}))) \\
&= (1-\alpha_i)^{-1} \left( 1 + \sum_{l \in N\setminus\{i\}} \alpha_l \right)^{-1} \\
&\quad \cdot \sum_{j \in N\setminus\{i\}} \left( 1 + \sum_{k \in N\setminus\{i,j\}} \alpha_k - \alpha_i \left( \sum_{l \in N\setminus\{i\}} \alpha_l \right) \right) (\bar{v}_j(f(a^{-i})) - \bar{v}_j(f(\bar{\boldsymbol{v}}))) \\
&\geq 0.
\end{aligned}$$

Here the first inequality follows from our assumption of nonnegative valuations and $c_i \geq 0$. The final inequality follows from the choice of $a^{-i}$.

Secondly, we show IR from Definition 3.7:

$$\begin{aligned}
u_i(v_i, \bar{\boldsymbol{v}}_{-i}) &= v_i(f(v_i, \bar{\boldsymbol{v}}_{-i})) + \alpha_i \sum_{j \in N\setminus\{i\}} \bar{v}_j(f(v_i, \bar{\boldsymbol{v}}_{-i})) - (1-\alpha_i)p_i(f(v_i, \bar{\boldsymbol{v}}_{-i})) \\
&= v_i(f(v_i, \bar{\boldsymbol{v}}_{-i})) + \alpha_i \sum_{j \in N\setminus\{i\}} \bar{v}_j(f(v_i, \bar{\boldsymbol{v}}_{-i})) - c_i \sum_{j \in N\setminus\{i\}} \bar{v}_j(f(a^{-i})) \\
&\quad - \sum_{j \in N\setminus\{i\}} \left( \frac{1 + \sum_{k \in N\setminus\{j\}} \alpha_k}{1 + \sum_{l \in N\setminus\{i\}} \alpha_l} - \alpha_i \right) (\bar{v}_j(f(a^{-i})) - \bar{v}_j(f(v_i, \bar{\boldsymbol{v}}_{-i}))) \\
&= v_i(f(v_i, \bar{\boldsymbol{v}}_{-i})) + (\alpha_i - c_i) \sum_{j \in N\setminus\{i\}} \bar{v}_j(f(a^{-i})) \\
&\quad + \sum_{j \in N\setminus\{i\}} \frac{1 + \sum_{k \in N\setminus\{j\}} \alpha_k}{1 + \sum_{l \in N\setminus\{i\}} \alpha_l} (\bar{v}_j(f(v_i, \bar{\boldsymbol{v}}_{-i})) - \bar{v}_j(f(a^{-i}))) \\
&\geq v_i(f(v_i, \bar{\boldsymbol{v}}_{-i})) \\
&\quad + \sum_{j \in N\setminus\{i\}} \frac{1 + \sum_{k \in N\setminus\{j\}} \alpha_k}{1 + \sum_{l \in N\setminus\{i\}} \alpha_l} (\bar{v}_j(f(v_i, \bar{\boldsymbol{v}}_{-i})) - \bar{v}_j(f(a^{-i}))) \\
&= v_i(f(a^{-i})) + \left(1 + \sum_{l \in N\setminus\{i\}} \alpha_l \right)^{-1} \\
&\quad \cdot \sum_{j \in N} \left(1 + \sum_{k \in N\setminus\{j\}} \alpha_k \right) (\bar{v}_j(f(v_i, \bar{\boldsymbol{v}}_{-i})) - \bar{v}_j(f(a^{-i}))) \\
&\geq 0
\end{aligned}$$

which proves the claim. The first inequality follows from the assumption of nonnegative valuation functions and from $c_i \in [0, \alpha_i]$. The final inequality follows again from the former assumption and from the choice of $f$. $\square$

# 6 Influence of altruism

In this chapter we use the models and mechanisms from Chapter 4 and 5 to make concrete the (positive or negative) influence altruism can have on results in mechanism design. First, we look into the effects altruism has on the properties of VCG mechanisms in the different models. We establish that in the player-oriented and welfare-oriented model, VCG mechanisms are no longer truthful. However, in the omnistic model they still are, and moreover the set of them satisfying NPT and IR under nonnegative valuations is larger than in standard mechanism design. Secondly, we look at the specific problem of financing a public project, which can be cast into the mechanism design setting and for which altruism turns out to have favorable effects in the omnistic model. Thirdly and lastly, we look into redistributing payments in the single-allocation setting, for which improved results in terms of budget-balancedness can be obtained in the welfare model.

## 6.1 VCG mechanisms

As was said in the Introduction, mechanism design deals with the problem of having egoistic players when we want to reach an objective for a larger group as a whole. When players are altruistic, and care more about this group, we would intuitively expect to be in a 'better' situation. However, when unknown or not accounted for, altruism might also have negative effects. Most importantly, it is not directly clear whether VCG mechanisms are still truthful when altruism is present, and therefore whether they can still be used effectively. It might be that a player is now ex-post motivated to lie about her valuation in order to improve the welfare of others, if this only comes at a small cost to her own egoistic utility. If this is indeed the case, we might sometimes prefer to have predictable egoistic players rather than players exhibiting some unknown altruism, or at least we should take care when to apply VCG mechanisms.

In this section we show that truthfulness of VCG mechanisms indeed no longer holds in the player utility and welfare model. However, in the omnistic model it remains valid. In the latter model altruism hence has only positive effects: the most popular type of mechanism in the literature can still be used, and on top of this new opportunities arise, as we already saw with the possibility to tweak $c_i$ for altruism-balanced mechanisms using the altruism-balanced payment rule. In this section we furthermore establish that VCG mechanisms with an extended form of the Clarke pivot rule allow the designer to extract extra payments from the players if necessary. In the next section, we will apply the resulting mechanisms to the public project problem.

We start by showing an interesting condition on a *model* in the utility framework for it to have VCG mechanisms satisfy the conditions of Proposition 5.2, and by those be truthful mechanisms in the model. It turns out that in such a case the altruistic feelings $g_i$ of a player $i$ cannot be influenced by her report

$\bar{v}_i$, not even indirectly through the alternative that is chosen by the mechanism.

**Proposition 6.1** For VCG mechanisms the conditions of Proposition 5.2 are satisfied if and only if applying such a mechanism makes that for some $h'_i : V_{-i} \to \Re$

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = h'_i(\bar{\boldsymbol{v}}_{-i}).$$

*Proof* Recall first that VCG payments are specified by:

$$p_i(\bar{\boldsymbol{v}}) = h_i(\bar{\boldsymbol{v}}_{-i}) - \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

Furthermore, since VCG maximizes social welfare we need:

$$\zeta_i(\bar{\boldsymbol{v}}) = \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}}))$$

Substituting these into condition (1) of Proposition 5.2 and rearranging gives:

$$g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) = h'_i(\bar{\boldsymbol{v}}_{-i})$$

as claimed. □

The omnistic model turns out to be special in that it has this specific feature, and thus VCG mechanisms are truthful in it:

**Theorem 6.1** When players value outcomes as specified by the omnistic model, all VCG mechanisms are truthful.

*Proof* We use Proposition 6.1:

$$\begin{aligned}
g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) &= \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) + \alpha_i p_i(\bar{\boldsymbol{v}}) \\
&= \alpha_i \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) + \alpha_i \left( h_i(\bar{\boldsymbol{v}}_{-i}) - \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) \right) \\
&= \alpha_i h_i(\bar{\boldsymbol{v}}_{-i})
\end{aligned}$$

where the first equality is by definition of $g_i$ for the omnistic model (Definition 4.3), and the second substitutes the VCG payments. □

Note on the side that as we did not need $\alpha_i \geq 0$ or $\alpha_i \leq 1$ anywhere in the above proofs, we may conclude that truthfulness is also conserved in situations with spiteful players and players that care more about others than about themselves.

Apart from showing how VCG mechanisms are truthful in the omnistic model, the proof of Theorem 5.1 gives a hint towards another nice feature: the designer can choose to extract extra payments from the players when they are altruistic. The Clarke pivot rule (which determines $h_i$) still makes sure that NPT is satisfied, as nothing changes in the payments with respect to the egoistic case. For the same reason, when valuations are nonnegative, individual rationality remains with respect to the standard utility the players experience. Because however the altruistic disposition $g_i$ gives each player some extra utility *proportional* to $h_i$, there are now more choices of $h_i$ that achieve NPT and IR:

**Definition 6.1 (Extended Clarke pivot rule)** We say that a VCG mechanism $(f, \boldsymbol{p})$ uses the extended Clarke pivot rule if for each player $i$ we have that

$$h_i(\bar{\boldsymbol{v}}_{-i}) = \left(1 + \frac{c_i}{1-\alpha_i}\right) \sum_{j \in N \setminus \{i\}} \bar{v}_j(a^{-i})$$

for some $c_i \in [0, \alpha_i]$ and $a^{-i} \in \arg\max_{a \in A} \sum_{j \in N \setminus \{i\}} v_j(a)$.

The Clarke pivot rule is obtained by setting $c_i = 0$ for all $i$. When player $i$ is spiteful ($\alpha_i < 0$) the set of feasible $c_i$'s is empty, indicating that in such a case we cannot obtain a VCG mechanism that satisfies both NPT and IR.
Also note that the designer only needs a (positive) lower bound on the altruism levels of the players in order to exploit them.

**Proposition 6.2** Whenever for all $i \in N$, $v_i \in V_i$ and $a \in A$ we have that $v_i(a) \geq 0$, a VCG mechanism that uses the extended Clarke pivot rule satisfies NPT and IR in the omnistic model.

*Proof* NPT follows directly from that it holds for the Clarke pivot rule ($c_i = 0$) and from the assumption of nonnegative valuation functions. IR follows similarly from that it holds for the Clarke pivot rule, and that a player $i$ pays an extra

$$\frac{c_i}{1-\alpha_i} \sum_{j \in N \setminus \{i\}} v_j(a^{-i})$$

which in her utility function $u_i^o$ is weakly more than compensated by (since $\alpha_i \geq c_i$):

$$\begin{aligned} g_i(\bar{\boldsymbol{v}}_{-i}(f(\bar{\boldsymbol{v}})), \boldsymbol{p}(\bar{\boldsymbol{v}})) &= \alpha_i h_i(\bar{\boldsymbol{v}}_{-i}) \\ &= \frac{\alpha_i}{1-c_i} \sum_{j \in N \setminus \{i\}} v_j(a^{-i}) \end{aligned}$$

where we use the proof of Theorem 5.1 to obtain the first equality. □

With these new possibilities established, let us see what they imply in a specific mechanism design setting.

## 6.2 Funding a public project

Imagine a group of people (e.g. a country) having the opportunity to undertake a project together at a commonly known cost $C$. When it is unclear how strongly each of the group members wants this project to be undertaken, it is hard to decide who will contribute to the investments that may need to be made. Moreover, in such a situation it is hard to decide whether it is worth it to undertake the project at all. The public project problem describes this setting. It can be formally defined as follows (after (Clarke, 1971)):

**Definition 6.2 (Public project problem)** The *public project problem* in mechanism design is characterized by the following conditions:

- The set of alternatives is given by $A = \{yes, no\}$.
- Player 1, whom we will call the *contractor* of the project, has a singleton valuation space $V_1 = \{v_e\}$, such that $v_e(yes) = -C$ for some $C \in \Re^*_+$ and $v_e(no) = 0$.
- Any player $i \neq 1$ has a valuation space $V_i$ such that for any $v_i \in V_i$ we have $v_i(yes) = w_i$ for some $w_i \in \Re_+$ and $v_i(no) = 0$.
- The design objective is the social welfare $SW$.

It is clear from the definition that all the designer needs from the players (except the contractor) is a report of their valuations $w_i$ for when the project is executed. We will refer to the report of player $i$ by $\bar{w}_i$. We further call the project *funded* if $\sum_{i\in N\setminus\{1\}} p_i(\bar{\boldsymbol{w}}_{-1}) \geq C$, so if enough payments are accumulated to by the players (excluding the contractor) to pay the contractor for executing the project.

This is a very practical situation, and it would be nice if the theory of mechanism design could provide us with a mechanism that makes sure that the project is undertaken when it should be undertaken, that is, when aggregate value created by the project is (weakly) larger than the necessary investments ($\sum_{i\in N\setminus\{1\}} w_i \geq C$). The latter is exactly what it means to maximize social welfare, so to achieve this we only need a truthful mechanism. Of course, we also need the group to be willing to pay for the investments when the project should be undertaken, for when the project cannot be funded knowing that it is worthy of funding is pointless.

Unfortunately, results in mechanism design show that there is no truthful mechanism that recovers all investment costs from the players excluding the contractor and that satisfies individual rationality for every such player (Nisan et al., 2007, Chapter 9). Even worse, there are no non-trivial instances at all in which the project costs are covered by a truthful mechanism. The positive result of this section is that in the omnistic model this impossibility result no longer holds. Here, we derive specific conditions for a project to be fully funded by the players (excluding the contractor).

Before we move on, let us clarify further the role of the contractor. Technically, she is a player. However, we do not need to pay her anything to report her valuation function truthfully, as her valuation space is a singleton set. She does not positively value the project for herself, so she cannot be asked to contribute to the investment costs. In fact, if everything works out, she should be the one to be paid by the rest of the players (for incurring the investment costs), so that funding the project corresponds to individual rationality for all players, including her. If she knows beforehand that she will not be compensated by the mechanism for executing the project, as is unfortunately the case with egoistic players, she has no reason to cooperate and execute the project when it should be executed. We shall assume that the investment costs are already adjusted for possible altruism of the contractor (so they are the 'real', final costs). In our treatment of the omnistic model we will hence take $\alpha_1 = 0$.

The fact that the contractor has a negative valuation $-C$ for when the project is undertaken means that our results for individual rationality for the VCG mechanism using the Clarke pivot rule do not (necessarily) extend to this situation. As already hinted at above, what happens is even worse: the project is funded in no non-trivial situation at all. Let us explain how these negative results in (standard) mechanism design come about.

Firstly, it is shown in (Nisan et al., 2007, Chapter 9) that given the individual rationality requirement and the choice of design objective, the VCG mechanism using the Clarke pivot rule is the only truthful mechanism that is feasible. To understand how this mechanism determines the payments in the case of the public project, it will turn out to be instructive to distinguish between two 'types' of players, dependent on their 'direct' role in whether the project is profitable or not. We will again refer to these types of players in discussing the results for when altruism is present, so let us define them explicitly:

**Definition 6.3 (Pivotal and non-pivotal players)** In an instance of the public project problem, after the reports are made, a player $i$ is termed *pivotal* whenever

$$\sum_{j \in N \setminus \{1\}} w_j \geq C$$

but

$$\sum_{j \in N \setminus \{1,i\}} w_j < C.$$

She is termed *non-pivotal* otherwise.

In other words, a pivotal player is essential to make the project profitable for the group as a whole. Note that by definition pivotal players only exist when the project is profitable, but that there do not need to be any pivotal players for a project to be profitable.
Let us prove the impossibility result:

**Proposition 6.3** When a VCG mechanism that uses the Clarke pivot rule is applied, the public project is funded ($\sum_{i \in N \setminus \{1\}} p_i(\bar{\boldsymbol{w}}_{-1}) \geq C$ if and only if at least one of the following two conditions holds:

- For some $i \in N \setminus \{1\}$ we have $\bar{w}_i \geq C$ and $\bar{w}_j = 0$ for all $j \in N \setminus \{1, i\}$.
- $\sum_{j \in N \setminus \{1\}} \bar{w}_j = C$.

*Proof* One can readily check that the VCG mechanism using the Clarke pivot rule makes each pivotal player $i$ pay exactly $C - \sum_{j \in N \setminus \{1,i\}} \bar{w}_j$, while a non-pivotal player never has to pay anything.

The if part for the first condition follows from that in this case every player $i$ with nonzero $\bar{w}_i$ is pivotal, so that

$$\begin{aligned}\sum_{i\in N\setminus\{1\}} p_i(\bar{\boldsymbol{w}}_{-1}) &= \sum_{i\in N\setminus\{1\}} (C - \sum_{j\in N\setminus\{1,i\}} \bar{w}_j) \\ &= (n-1)C - (n-2) \sum_{i\in N\setminus\{1\}} \bar{w}_i \\ &= C.\end{aligned}$$

In the situation of the second condition, secondly, the pivotal player $i$ trivially pays $C - 0 = C$.
Then the only-if part: when both conditions do not hold, take any pivotal player $i$. Note that without her, payments are $0 < C$, so she must exist for the project to be funded. She pays $C - \sum_{j\in N\setminus\{1,i\}} \bar{w}_j$, so in order to have the project funded, all other players must pay at least $\sum_{j\in N\setminus\{1,i\}} \bar{w}_j$. This is larger than zero, because the first condition does not hold. Thus, on average every player has to pay her own valuation. However, the only players that pay are pivotal players, and they pay their own valuation at maximum, because by definition of a pivotal player $i$ we have that

$$\begin{aligned}C - \sum_{j\in N\setminus\{1,i\}} \bar{w}_j &= C - \sum_{j\in N\setminus\{1\}} \bar{w}_j + \bar{w}_i \\ &\leq \bar{w}_i.\end{aligned}$$

So the only way to now fund the project is when all players with nonzero valuations pay exactly their own valuation, and are hence pivotal. By the above this happens only when $C - \sum_{j\in N\setminus\{1\}} \bar{w}_j = 0$. But then the second condition holds, which is in contradiction with our initial assumption. □

So a public project is only financed when there is only one person who benefits from it, or when there is no net benefit at all but just a break even between the value created and the investment costs incurred. Even worse, one can easily see that generally the larger the net benefits for the group from the project, and the more these are spread among the group members, the less payments the group can bring together to fund it. In the extreme, if no player at all is pivotal, zero payments are collected. This fact, that non-pivotal players pay nothing, is intuitive if one considers that given the reports of the other players, a non-pivotal player $i$ can still enjoy her full valuation $w_i$ for the project if she reports $\bar{w}_i = 0$. Non-pivotal players can, as a consequence, 'ride along' on their pivotal colleagues' high valuations for the project.

We now turn to the question how altruism can help here. More specifically, we look at the results that a VCG mechanism using the extended Clarke pivot rule with $c_i = \alpha_i$ for all $i$ can obtain in the omnistic model. We see that the impossibility result is overcome, and that now, even for relatively low levels of altruism, the project can in many cases be funded, as the example below illustrates.

**Example 6.1** Consider a public project setting in the omnistic model with $n = 11$ and $C = 10$, and in which all players are 0.2-altruistic. The VCG mechanism using the extended Clarke pivot rule with $c_i = \alpha_i$ for all $i$ is now applied, yielding that every player $i \in N \setminus \{1, 11\}$ reports a valuation $\bar{w}_i = 1$ and Player 11 reports a valuation $w_i = 5$. The mechanism prescribes that the project should be executed (as $\sum_{i \in N \setminus \{1\}} \bar{w}_i = 15 \geq C = 10$). Furthermore, it is readily checked that all 10 non-contractor players are required to pay exactly 1, which means that the project is funded. If we would have used the VCG mechanism using the 'standard' Clarke pivot rule instead, *total* payments would have amounted to 1, and thereby the project could have never been executed.

Example 6.1 already shows that the role division in paying between pivotal and non-pivotal players changes when altruism is exploited in the omnistic model. This might even happen to the extent that non-pivotal players pay more than pivotal players. In fact, the only players that pay more as a consequence of the designer exploiting altruism are non-pivotal players; pivotal players keep paying exactly the same amount. To best capture the exact influence of altruism on the funding of a project, we hence assume in the proposition below that there are only non-pivotal players. Such a situation is likely to occur when the group of players is large and when there are large net benefits to be gained from the project.

**Proposition 6.4** When all players are non-pivotal, the VCG mechanism using the extended Clarke pivot rule with $c_i = \alpha_i$ for all $i$ results in the funding of a profitable ($\sum_{i \in N \setminus \{1\}} \bar{w}_i \geq C$) public project in the omnistic model if and only if

$$\sum_{i \in N \setminus \{1\}} \frac{\alpha_i}{1-\alpha_i} \sum_{j \in N \setminus \{1\}} (\bar{w}_j - C) \geq C + \sum_{j \in N \setminus \{1\}} \frac{\alpha_j}{1-\alpha_j} \bar{w}_j.$$

*Proof.* Writing out the total payments for the non-contractor players yields:

$$\begin{aligned}
\sum_{i \in N \setminus \{1\}} p_i(\bar{\boldsymbol{v}}) &= \sum_{i \in N \setminus \{1\}} \left( \left(1 + \frac{\alpha_i}{1-\alpha_i}\right) \sum_{j \in N \setminus \{i\}} \bar{v}_j(a^{-i}) - \sum_{j \in N \setminus \{i\}} \bar{v}_j(f(\bar{\boldsymbol{v}})) \right) \\
&= \sum_{i \in N \setminus \{1\}} \left( \frac{\alpha_i}{1-\alpha_i} \left( \sum_{j \in N \setminus \{1,i\}} \bar{w}_j - C \right) \right) \\
&= \sum_{i \in N \setminus \{1\}} \frac{\alpha_i}{1-\alpha_i} \left( \sum_{j \in N \setminus \{1\}} \bar{w}_j - C \right) - \sum_{j \in N \setminus \{1\}} \frac{\alpha_j}{1-\alpha_j} (\bar{w}_j).
\end{aligned}$$

The second equality follows from that for non-pivotal players we have $a^{-i} = f(\bar{\boldsymbol{v}}) = yes$. As the total payments need to be larger than $C$ to fund the project, the above proves the claim. $\square$

What does Proposition 6.4 tell us? We see first of all that the project is more likely to be fully funded when it is more profitable for the group to undertake it. This relation is rather satisfying, especially if you compare it with the standard

'egoistic' case. Recall that there we noted that, paradoxically, the larger the net benefits of the project are, the less likely it is that it will be financed.
Secondly, even though it has the largest effect when concentrated among the players that have the lowest valuations, we observe that altruism always has a positive effect on the likelihood of funding the project. Because all players are non-pivotal, $\bar{w}_i \leq \sum_{j \in N \setminus \{1\}} \bar{w}_j - C$ for all $i$, and hence when $\alpha_i$ increases for any $i$, any increase in the right-hand term is always superseded by an increase on the left-hand side.

If we consider universal altruism ($\alpha_i = \alpha$ for all $i$) and slightly rearrange the inequality, the result is even easier to interpret:

**Corollary 6.1** When all players are non-pivotal and $\alpha_i = \alpha$ for all $i$, the VCG mechanism using the extended Clarke pivot rule with $c_i = \alpha$ for all $i$ results in the funding of a profitable public project in the omnistic model if and only if

$$\alpha \left( \sum_{j \in N} \bar{w}_j - C \right) \geq \frac{C}{n-1}.$$

Here we can even more clearly see the positive dependence on both the altruism levels and the project profits. Moreover, it becomes clear that the number of players $n$ is also positively related to the likelihood of financing the project. This last thing is explained first of all by the fact that every added player by assumption is altruistic at level $\alpha$, and second of all by the fact that, all other things equal, when the number of players increases each player on average becomes 'more' non-pivotal, in the sense that her valuation is smaller relative to the gap $\sum_{j \in N} \bar{w}_j - C$. As noted before, only non-pivotal players contribute extra due to altruism, and the 'more' non-pivotal they are, the more they contribute.

As it is only because of this altruism of non-pivotal players that budget-balancedness can be attained, one might expect that the above case (of only non-pivotal players) is somewhat of a best case scenario for funding the project, which would weaken its implications. In fact, the opposite is true: given total net benefits ($\sum_{j \in N} \bar{w}_j - C$) having only non-pivotal players is in most cases a worst case scenario, as the example below illustrates.

**Example 6.2** Consider a public project setting in the omnistic model with $n = 4$, $C = 1$ and semi-altruistic players ($\alpha_2 = \alpha_3 = \alpha_4 = 0.5$). The VCG mechanism using the extended Clarke pivot rule is applied. We compare two possible report vectors $\bar{\boldsymbol{w}}$ and $\bar{\boldsymbol{z}}$ with equal aggregate valuations $\sum_{j \in N \setminus \{1\}} \bar{w}_j$ and $\sum_{j \in N \setminus \{1\}} z_j$, and hence with equal total net benefits. In the first situation we have only non-pivotal players: $\bar{w}_2 = \bar{w}_3 = \bar{w}_4 = 0.6$. In the second we have one pivotal and two non-pivotal players: $\bar{z}_2 = 1$ and $\bar{z}_2 = \bar{z}_3 = 0.4$.
When players would have been egoistic, the first situation would have yielded a total payment of 0 and the second only 0.2. Because of the presence of altruism, however, the first situation now yields a payment of 0.6 and the second a payment

of 1. Thus, in this second situation the project can be financed. In fact, it can be readily seen that when only non-pivotal players are present total investments here always amount to 0.6, whereas when players are of mixed types they may amount to anything between 0.6 and 2.6.

It turns out that the project is most likely to be financed if altruism is concentrated within the group of non-pivotal players, and if players are either 'very' pivotal or 'very' non-pivotal. Proposition 6.4 can therefore indeed be seen as a sort of worst case (instead of best case) boundary condition for the project to be financed, illustrating the rather large beneficial influence altruism can have in funding a public project in the omnistic model.

## 6.3 Single-item allocation and redistribution

In this section, we study the influence of altruism in another specific setting: that of the *single-item allocation problem* (also known as the *single-item auction*). We start by formally defining this setting:

**Definition 6.4 (Single-item allocation problem)** The *single-item allocation problem* in mechanism design is characterized by the following conditions:

- The set of alternatives is given by $A = \{a_1, ..., a_n\}$, where $a_i$ denotes the case in which player $i$ obtains the item that was to be allocated.
- Each player $i \in N$ has a valuation space $V_i$ such that for any $v_i \in V_i$ we have $v_i(a_i) = w_i$ for some $w_i \in \Re_+$ and $v_i(a_j) = 0$ for any $j \neq i$.
- The design objective is the social welfare $SW$.

Note that maximizing social welfare here corresponds to allocating the item to a player with the highest reported valuation for it, that is to an $i \in \arg\max_{j \in N} \bar{w}_j$.

The single-item allocation problem provides an intuitive demonstration of the necessity and use of mechanism design. When all players are egoistic, and no payments would be issued, it is clear that it is best for each player to report a valuation as high as possible. If players would indeed act in this way, no information at all would be extracted from the reports. The Vickrey or second-price auction (Vickrey, 1961) for this problem is the standard example of a mechanism in any introduction to mechanism design, and is an instance of applying a VCG mechanism using the Clarke pivot rule. In it, the object under auction is given to the highest bidder, who pays the second-highest bid (the rest of the players pays nothing). It is readily checked that this indeed is a VCG mechanism, and that it (hence) warrants truthfulness.

In some contexts, the allocator/auctioneer in question might want to extract as many payments as possible. However, when there is no clear destination for the payments made for the object to be allocated (e.g. when the object has no previous owner), the designer might want to allocate the item in the social welfare optimizing way with as few payments extracted as possible. This latter case is the one we treat here, and has received quite some recent interest in the literature (see for instance Cavallo (2006), Apt et al. (2008) and Guo & Conitzer (2010)). To keep the favorable properties of the second-price auction (i.e. NPT and IR) but still generate fewer net payments than the second highest bid, there has been a search for ways to redistribute the generated payments by means of *redistribution mechanisms* (Cavallo, 2006). Obviously, these redistribution mechanisms should not have strategic implications, for otherwise truthfulness would be lost, and we might allocate the item to the 'wrong' bidder after all. For instance, the naive option of equally dividing the payments over the players may result in such situations. When it is applied, the second-highest bidder has incentives to increase his bid up to the level of the highest bidder, since the former receives a share of the payments the latter makes. In fact, it is shown in (Cavallo, 2006) that there is no redistribution mechanism that conserves truthfulness, IR and NPT and redistributes the full payments. Finding the best redistribution mechanism (in that it redistributes most of the payments) is hence a non-trivial task. Apt et al. (2008) provide a characterization of a class of mechanisms that is optimal with respect to the aggregate utility it generates. Bailey (1997) and Cavallo (2006) came up with an intuitive mechanism which turns out to be in this class. Let us define this mechanism, as we will use and slightly adjust it in our study of the effects that altruism can have here. We will order the players by the size of their reported valuations (breaking ties arbitrarily), so that $\bar{w}_i$ corresponds to the $i$th highest report.

**Definition 6.5 (Bailey-Cavallo redistribution mechanism)** The *Bailey-Cavallo redistribution mechanism* (BCR mechanism) for the single-item allocation problem works as follows:

- First, the second-price auction (VCG mechanism using the Clarke pivot rule) is applied.
- Second, the collected payments $\bar{w}_2$ are partly redistributed in the following way: Player 1 and 2 (the 'winner' and 'runner-up') both receive $\frac{\bar{w}_3}{n}$ and Player 3 up to $n$ all receive $\frac{\bar{w}_2}{n}$.

Note first that the BCR mechanism runs no deficit, but generally does not redistribute all payments: the net payments to the mechanism are exactly $\frac{\bar{w}_2 - \bar{w}_3}{n} \geq 0$. Still, this is a huge improvement to the $\bar{w}_2$ that is obtained by just applying the second-price auction.
Secondly, when looked at from the right perspective, it is not hard to see that the BCR mechanism is indeed truthful. Every player is paid a fraction of $\frac{1}{n}$ of the second-highest bid when only considering bids other than her own. As no player can have any influence on the size of the redistribution payment she receives by reporting differently herself, the truthfulness of the second-price auction is conserved.

Let us now see how altruism may be of influence here. We consider the welfare-oriented model, for it seems to fit the setting: apart from wanting to receive the object herself, each player may to a certain extent like the object to end up with the right person, which the term $g_i$ in the welfare model represents. The attentive reader may now object that the welfare model indeed works well for the single-allocation setting, but that when redistribution is made the designer attains exactly the role of 'mediator' that we warned for when applying the welfare-oriented model (see Section 4.3). However, we argue that the model can still be applied, for the redistribution part of the *altruism-adjusted BCR mechanism* that we will propose has no influence on the incentives of the players. Technically then, the designer indeed functions as a 'mediator' here, but in establishing truthfulness and individual rationality (which is what the model is used for) the redistribution part of the mechanism is not involved. We will apply the AAVCG mechanism using the altruism-adjusted Clarke pivot rule for the first, 'non-redistribution' part of the mechanism, and there these considerations do count, but now only positive payments to the mechanism are made.

The mechanism we propose is as follows:

**Definition 6.6 (Altruism-adjusted BCR mechanism)** The *altruism-adjusted Bailey-Cavallo redistribution mechanism* (AABCR mechanism) for the single-item allocation problem works as follows:

- First, the AAVCG mechanism using the altruism-adjusted Clarke pivot rule with $c_i = 0$ for all $i$ is applied.
- Second, the collected payments $(1 - \alpha_1)\bar{w}_2$ are partly redistributed in the following way: Player 1 and 2 (the 'winner' and 'runner-up') both receive $(1 - \alpha_1)\frac{\bar{w}_3}{n}$ and Player 3 up to $n$ all receive $(1 - \alpha_1)\frac{\bar{w}_2}{n}$.

The net payments made to the AABCR mechanism then amount to $(1 - \alpha_1)\frac{\bar{w}_2 - \bar{w}_3}{n}$, an improvement to the regular BCR mechanism whenever the winner of the object is altruistic.

Three remarks should be made. Firstly, note that the redistribution indeed still has no effects on truthfulness: it does not have any impact on the altruistic disposition term $g_i$ of a player, and there is still no influence the player can exert (by reporting untruthfully) to change the size of her own redistribution payment. Secondly, note that the altruistic disposition that is of influence here is that of Player 1, the player reporting to value the object the most. This stands in interesting contrast to the results of the previous section on the public project problem, in which influence of altruism was the largest through players that valued the project the *least* (those that were 'most' non-pivotal). And lastly, note that a positive effect proportional to the altruism level is even obtained *without* redistributing the payments made: the AAVCG mechanism we use only makes the winner pay $1-\alpha_1$ times the second-highest bid.

We may conclude that also for countering the problem of high net payments in a single-item allocation setting in the welfare-oriented model, altruism may be of significant beneficial influence.

# 7 Conclusion

In this work, we have provided methods to model altruistic preferences in mechanism design, and have studied the changes these preferences make to established results in standard mechanism design.

We started by introducing the utility framework to model altruism, and by providing intuitive models within this framework. In our comparison with models that had been used so far in the literature, we found that depending on the situation under study some models were preferable to others. In particular, we noted that player-oriented model adopted by most other researchers best fits situations in which players only care about other *players*, and not about others in general. The welfare-oriented model that we introduced is intuitive when players do not necessarily care about others personally, but only about the total welfare that is created. With this model, however, care should be taken when the mechanism functions as a mediator when it comes to payments. Our omnistic model, on the other hand, fits any situation in which players are truly altruistic in a non-directed way (omnistic), in the sense that they care about others equally. It can be easily adjusted to the directed case as well; the player-oriented model then becomes a special case of it. Lastly, we noted that the regret-based model proposed by Cavallo provides an interesting alternative to the utility framework, but has the downside of not considering the *extent* to which certain outcomes may benefit others when modeling altruism for a specific player.

Secondly, we turned to the design of mechanisms for when altruistic players are present. We characterized truthful mechanisms for the general utility framework, and derived a template to find truthful mechanisms with the no positive transfers and individual rationality properties if valuations are nonnegative. We then applied our findings to the two new models we introduced and provided such mechanisms, which included the well-known VCG mechanisms and Clarke pivot rule as special cases.

With our new mechanisms in hand, we were ready to study the influence of altruism on the results of applying mechanisms. We did this firstly by examining whether (unknown) altruism may have negative effects in the sense that our standard VCG mechanisms are no longer truthful. For the welfare-oriented model and player-utility model we observed that this is indeed the case, but for our most-preferred omnistic model we found that VCG mechanisms remain truthful. Moreover, when altruism levels are known to some extent, an extended version of the Clarke pivot rule provides us with a leeway in choosing payments while retaining truthfulness, NPT and IR.
As a second investigation into the influence of altruism, we studied the public project problem and found that the impossibility results with respect to funding no longer hold in the omnistic model. In fact, we derived worst-case boundary conditions that show that even for moderate levels of altruism a profitable

project can be funded in most cases.
Lastly, we looked into the problem of minimizing net payments in a single-item allocation setting in the welfare model. Here we found a reduction in total payments needed proportional to the altruism level of the winner, which resulted mainly from the fact that part of her incentives to report truthfully were already provided by her altruistic disposition.

Putting all our results together, we may conclude that even when players are only partially altruistic, this has a mostly beneficial impact with regards to a designer being able to solve the problems posed in mechanism design. Of course, for this impact to indeed be positive, the right tools (mechanisms) need to be applied, such as those we have provided here.

We finish by providing some recommendations for future research in this area.
Firstly, to gain further insight also into other types of other-regarding behavior than altruism, it would be useful to obtain a full characterization of models $g_i$ for which truthful mechanisms exist.
Secondly, note that in most of this work we have assumed that altruism levels are known to the designer. We saw that in the omnistic model, when this is not or only partly the case, VCG mechanisms retain their truthfulness, but our other mechanisms are mostly based on knowledge of the altruism levels. When they are unknown, or when there is only partial (e.g. probabilistic) information on these, it is an open question whether truthful mechanisms can be designed or approximated that still exploit the altruism that is present. For instance, one could imagine a mechanism in which apart from the valuation functions players also report an altruism parameter that fits a given model. Developing such a multi-parameter setting or proving impossibility results there is certainly a challenging and possibly a rewarding direction of research.
Lastly, the gap between human behavior and game theory may be further bridged by developing models of altruism using experimental results from behavioral economics. In a similar way as Cavallo restricts utility functions in his regret-based approach, utility functions may be restricted to those that agree with empirical results. The resulting models can then be subjected to further mathematical analysis to obtain predictions that may again be tested by behavioral economists. In this way, the two disciplines may converge and together provide mote insight into the workings and implications of human (ir)rationality.

# 8 References


Andreoni J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. In *Econometrica* 70(2) (pp. 737-753).

Apt, K., Conitzer, V., Guo, M., & Markakis, E. (2008). Welfare undominated Groves mechanisms. In *Internet and Network Economics* (pp. 426-437). Springer Berlin Heidelberg.

Apt, K. R., & Schäfer, G. (2014). Selfishness level of strategic games. In *Journal of Artificial Intelligence Research* 49 (pp. 207-240).

Bailey, M. J. (1997). The demand revealing process: to distribute the surplus. In *Public Choice* 91(2) (pp. 107-126).

Brandt, F., & Weiss, G. (2002). Antisocial agents and Vickrey auctions. In *Intelligent agents VIII* (pp. 335-347). Springer Berlin Heidelberg.

Cavallo, R. (2006). Optimal decision-making with minimal waste: Strategyproof redistribution of VCG payments. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems* (pp. 882-889). ACM.

Cavallo, R. (2012). Efficient Auctions with Altruism. In submission. Published online at http://www.eecs.harvard.edu/ cavallo/papers/cavallo-altru.pdf.

Caragiannis, I., Kaklamanis, C., Kanellopoulos, P., Kyropoulou, M., & Papaioannou, E. (2010). The impact of altruism on the efficiency of atomic congestion games. In *Trustworthly Global Computing* (pp. 172-188). Springer Berlin Heidelberg.

Charnes G., & Rabin, M. (2002). Understanding social preferences with simple tests. In *Quarterly Journal of Economics* 117 (pp. 817-869)

Clarke, E. (1971). Multipart pricing of public goods. In *Public Choice* 8 (pp. 19-33).

De Marco, G., & Morgan, J. (2011). Altruistic Behavior and Correlated Equilibrium Selection. In *International Game Theory Review*, 13(04), (pp. 363-381).

Groves, T. (1973). Incentives in teams. In *Econometrica*, 41 (pp. 617-631)

Guo, M., & Conitzer, V. (2010). Optimal-in-expectation redistribution mechanisms. In *Artificial Intelligence*, 174(5), 363-381.

Kahneman, D. (2011). *Thinking, fast and slow.* New York: FSG.

Kucuksenel, S. (2012). Behavioral mechanism design. In *Journal of Public Economic Theory*, 14(5) (pp. 767-789).

Nisan, N., Roughgarden, T., Tardos, E. & Vazirani, V. (Eds.). (2007). *Algorithmic game theory.* New York: Cambridge University Press.

Tang, P., & Sandholm, T. (2012). *Optimal Auctions for Spiteful Bidders.* In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence.*

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. In *Journal of finance*, 16 (pp. 8-37).