# COMPENSATION SCHEMES IN COLLECTIVE DECISION MAKING

Memoria presentada por Juan Perote Peña para optar al
grado de doctor en Ciencias Económicas
Departament d'Economia i d'Història Econòmica,
International Doctorate in Economic Analysis (IDEA)
Universitat Autònoma de Barcelona (UAB)
Dirigida por el Doctor Salvador Barberà Sández

Bellaterra, Diciembre 2000

A MIS PADRES MARIA PILAR Y JUAN

# CONTENTS

Acknowledgments. 3

1. Introduction. 4

2. Solidarity in terms of reciprocity. 7

3. A Social Choice trade-off between alternative fairness concepts: solidarity versus rigidity. 34

4. Dominant strategies implementation when compensations are allowed: a characterization. 60

5. Dominant strategies implementation of the Critical Path Allocation in the project planning problem. 96

# ACKNOWLEDGMENTS

Major acknowledgment goes to Salvador Barberà, my thesis advisor, for his support and guidance. It has been a great pleasure to be his student.

I thank my brother Javier and my friend and co-author Paco Candel for the encouragement they gave to me while pursuing their own doctoral research. Their enthusiastic view of Economics has always been inspiring.

I am very grateful to Alfredo Pastor for the support and advice I received from him during the years I was working in IESE. The present memory would not exist without him.

I am indebted to José Alcalde, Carmen Beviá, Luis Corchón, Bashkar Dutta, Theodore Groves, Jordi Massó, Hervé Moulin and Arunava Sen for their interest in my work and friendly advice at various points in the process.

I wish to thank everybody in the *Departament d'Economia i d'Història Econòmica,* specially to my lecturers in the IDEA program, from whom I received a solid background in economic theory.

Acknowledgments extend to many people in the *School of Management and Economics at The Queen's University of Belfast* for making it easy for me to come to Barcelona and to continue my research.

The research leading to this thesis has been partially supported by a *Beca de Formación del Personal Investigador* from the *Ministerio de Educación y Ciencia* of the Spanish Government.

Last, but not least, I acknowledge my parents Maria Pilar and Juan, my girlfriend Mercedes and my friends in Zaragoza, Barcelona, Murcia and Belfast for their restless support even in the hardest moments.

# 1 Introduction

This memory consists of four independent essays on Social Choice Theory and Mechanism Design. The first two chapters are particularly related since they share the common aim of relaxing an idea of the concept of *solidarity* called *replacement monotonicity* or welfare-domination under preference-replacement proposed in the literature by Moulin (1987) and Thomson (1993, 1995a, 1995b).[1] This property requires that whenever any individual changes his preferences and this shifts the social choice, every other agent should move in the same welfare direction, i.e., either all of them gain or all of them lose with the change. Chapters 2 and 3 share the same basic model of public good provision in which no monetary compensations are allowed. Agents are assumed to have single-peaked preferences defined on a single unidimensional good, and we try to investigate and overcome the limited number of Social Choice procedures that are replacement monotonic within this setup.

*Chapter 2* introduces a new property of solidarity in terms of *reciprocity* for these environments. This property allows for a richer class of social choice procedures than Thomson's solidarity concept of welfare-domination under preference-replacement. Our proposal reflects somehow an introspective concept of solidarity. It requires that when somebody changes his preferences and the social decision is altered, all the remaining agents can be sure that the same agent who has actually changed would have been affected in a reciprocal way if they had changed likewise. Characterizations of the rules satisfying reciprocity in both a strong and a weak version are therefore provided.

*Chapter 3* relaxes replacement monotonicity in a more direct way, understanding solidarity in the sense of the proportion of society gaining or losing together when somebody changes his preferences. It is shown that solidarity is in direct conflict with other fairness criteria. Actually, achieving *centrist* or equitable outcomes leads to the minimal possible solidarity degree. In Chapter 3, we propose reasonable measures of both concepts of *solidarity* and *rigidity* of social choice procedures and use them to prove the existence of such a trade-off. Characterizations of those rules with the highest levels of solidarity and rigidity are offered and a menu of voting procedures is proposed and classified for consideration when society agree some feasible

[1] The references in this introduction can be found in the corresponding chapters.

solidarity-rigidity level to hold.

Chapters 4 and 5 are concerned with the problem of finding interesting strong incentive compatible mechanisms in two different production contexts when compensations to induce truthful behavior are allowed, so we tackle a problem of implementation in dominant strategies.

In *Chapter 4* we deal with an implementation problem where the compensation scheme is not part of the feasible alternatives, but discretional to the planner. The incentive compatible compensation mechanisms we characterize make complete sense in production economies in which, for example, some divisions within a firm have to share some input available in a fixed amount or have to decide the level of some public input or service and information about the individual technologies is private. Within this framework, we assume that the planner or principal have the ability to partially design the functional form of the divisions' final payoffs to incentive their truthful behavior in the contract offered to the divisions.

There are one important difference between our approach to the implementation problem in Chapter 4 and the mainstream in the implementation literature: the treatment of the concept of *alternatives* or social states. Most of the literature implicitly define alternatives as *everything that can be allocated and is not fixed or given by nature*, and since the agents' *types* are the private information given by nature, the type determines the final payoff the agent receives for each alternative, so any variable feature that can affect any agent's final payoff given her type will be an alternative, and they are the only way in which the planner can modify the agents' objective functions. Besides, the planner's objectives are defined on the set of alternatives for any type of the agents. In this sense, our alternatives should include some set of functional forms of the agents' final payoffs with the additional restriction on the planner's preferences that she does not care about these functional forms. Our approach, instead, define alternatives as *any variable feature that can affect the planner's objectives given the agents' types* and, since the planner is indifferent about the functional form of the agents' objectives for all types, the set of relevant alternatives is reduced to a smaller set. This framework implies that the new set of alternatives is no longer the only way an agent's objective function or final payoff can vary given his type: the planner can actually affect the agents' payoffs in two ways: changing the alternative chosen -which can change her own preferences too- and directly shifting the agents' payoffs within some admissible bounds -leaving the planner unaffected-.

*Groves' mechanisms* -see Groves (1973, 1975) and Green and Laffont (1977)- for providing public goods or public inputs can be interpreted as a restricted case of our framework.

In *Chapter 5* we deal with the same problem of designing a contract between the members of a complex productive organization -say the divisions within a firm- to give incentives for truthful behavior in the presence of private information about parts of the technologies, but now there is a more important conflict of interests between the planner and the divisions. In particular, we propose a simple task allocation model known in the network programming literature as the *Critical Path Method*, CPM or PERT. Divisions have to undertake time-consuming tasks that have to be allocated in a network. The principal would like all tasks to be finished as soon as possible, but informed divisions are all interested in delaying their own tasks if they are not conveniently compensated. We prove the existence in this environment of anonymous and strategy-proof mechanisms that are efficient in the sense of minimizing the total amount of time employed in finalizing the project and can also be balanced, so full efficiency is achieved. Moreover, an impossibility result emerges if we require them further to be individually rational in the sense of guaranteeing a minimum net payoff for every agent in every circumstance.

Since the four chapters are quite independent, we have decided for the sake of clarity of exposition to include the relevant bibliography and the main conclusions and comments at the end of each one.

# 2 SOLIDARITY IN TERMS OF RECIPROCITY

## 2.1 Introduction

A solidarity principle applying to the fair allocation problem was introduced by Thomson [12] under the name of *replacement principle*. The idea is the following: every allocation problem can be described by some parameters or *data*, such as the set of agents involved in the decision, the description of their preferences or the possible amount of resources and their distribution among the agents. The replacement principle imposes solidarity among agents in the following sense. If some component of the *data* changes its value within the admissible domain, every agent should be affected in the same direction: either all of them improve their position or all of them lose. It is argued that *fair* and acceptable social choice rules should fulfill this equal treatment property when facing exogenously given shocks. The replacement principle has been widely explored in the literature in different contexts. When we consider the population as the *relevant* variable parameter, Thomson's [15], [16] concept of *population monotonicity* is the accurate translation of the replacement principle: every agent should lose when we add new agents to those initially present, since the growth of the population can be seen as a restriction of the opportunities available to society. This property was investigated by Moulin [7] in connection with strategy-proofness and by Thomson [11], [15], [16] and Ching and Thomson [4] in the context of single-peaked preferences. If we focus on a change on the amount of available resources, then, the replacement principle takes the form of the *resource monotonicity*, a property analyzed by Thomson [14].

The specific version of the replacement principle we are going to discuss here applies when the preferences of some individuals change. It was first defined by Moulin [8] under the name of *replacement domination* and later by many authors with the names replacement monotonicity or *welfare-domination under preference-replacement* ($WDUPR$). It requires that if somebody changes his preferences, and this shifts the social decision affecting the remaining agents, then, they should all move in the same direction: they should either all gain or all lose after the change. This property has been analyzed in the two contexts of private and public goods economies in

Thomson [17], [18] and [13] respectively -see Thomson [19] for a comprehensive survey-.

We will consider here the provision of a public good, where there are a continuum of alternatives described by an interval of the real line. It has been shown by Thomson [13] that with single-peaked preferences, the only replacement monotonic and efficient social choice functions are those functions that choose a fixed given point in the interval if it is Pareto optimal and choose the nearest efficient point to this one if it is not.

The intuition is easy: let us first remember that a preference relation is single-peaked if there is a best preferred alternative and the further is an alternative from this ideal, the worse it becomes. We can, then, imagine a preference profile in which every individual allocates his best point at one of the extremes of the interval, in such a way that all the people's ideals are polarized at both extremes. Suppose that the chosen alternative in this situation is, for example, $a$. Let us think now of any other preference profile; we can construct a sequence of profiles just by iteratively changing each agent's preferences in the first profile to the preferences in the second. Since some individual will be best at each of both extremes in all the intermediate profiles -except perhaps in the last one-, the selected alternative in these profiles should be $a$ to preserve replacement monotonicity, because otherwise any change along the sequence would affect the others differently. The only shift allowed can occur if $a$ is not efficient -$a$ is not contained in the interval defined between the lowest and highest peak of preferences- and in this case, it can be proved that efficiency requires to shift the chosen point from $a$ to the nearest efficient point to $a$, because any other choice will violate either efficiency or $WDUPR$.

The above class of social choice functions constitutes a subclass of the family of *Generalized Condorcet winner solutions* defined by Moulin [6], and we feel that they are in fact very far from desirable. They are quite trivial decision rules. They weight excessively an arbitrary *status quo*, so they are very insensitive to individual preferences than other more flexible procedures one can think of.

The present chapter starts from this criticism and at the same time tries to provide a reasonable alternative to Thomson's principle of $WDUPR$. In order to enlarge the class of procedures, a new and intuitive concept of solidarity among agents -we call it *reciprocity*- is introduced in both a strong and a weak version. The social choice functions and voting schemes that preserve both reciprocity and anonymity are fully characterized.

The alternative property proposed here tries to embody part of the substance of the original idea of $WDUPR$, but differs from this in the sense that reciprocity can be considered as a somehow introspective conception of solidarity. Let us think of the society just before deciding what social choice function -from now on SCF- is going to be used in choosing the level of some public good. People would like to accept a procedure that embodies some idea of solidarity in the sense that this rule provides some form of protection for every individual against the possible shifts in choice caused by the changes of preferences of others.

Thomson's requirement of $WDUPR$ can be reinterpreted in this context. The ex-ante social contract in the SCF guarantees that if any individual changes his preferences, the new value of the function is such that everybody moves in the same direction -all of them gain or all of them lose with the change-.

Although $WDUPR$ is a useful property, it turns out to be so strong that only quasi-trivial and conservative SCFs are allowed. Our reciprocity condition can be seen as another type of insurance: agents are no longer treated equally than the rest of individuals who maintain their original preferences, but equally than the agent who changed his own. The idea is as follows: people may now gain or lose when somebody changes, but if I lose, I want to be sure that the agent who has caused my loss would be in the same situation than me if I had changed likewise and shifted the social decision. He would have been moved by my change in the same direction than I was moved by him.

By considering such contracts before deciding the optimal rule for society, people might be ready to accept this weaker and introspective concept of solidarity. Moreover, its philosophy is very intuitive and can be heard in the real world -people usually are much more permissive with the impositions of others when they dislike them if they know that they would be treated equally under similar situations-. The proposed property -in its weak version- allows for a larger and more flexible class of functions than those allowed by Thomson's $WDUPR$, although Thomson's class of efficient, replacement monotonic SCFs satisfies reciprocity.

The structure of this chapter is as follows. We first introduce the model in *Section 2*. In *Section 3*, the reciprocity properties are proposed and results are presented. It is shown that Thomson's class is a narrow subset of our class. We close with some comments and conclusions.

## 2.2 The model

Consider a society defined by a set of *agents or individuals*: $N = \{1, ..., n\}$, indexed by $i$ and sometimes by $j, h$ and $l$. Society must make decisions from some predetermined set of mutually exclusive *alternatives,* represented by $A$, whose elements will be denoted by $x, y, ... \in A$. The set of alternatives will sometimes be finite -representing discrete levels of the provision of some public good- and sometimes infinite: a closed interval of the real line, normalized for simplicity to the interval $[0, M]$ and standing for the continuous amount of the public good or the location of some public utility.

Every individual $i$ is endowed with a complete *preference relation* over the set of alternatives denoted as $R_i$ from some set of possible preferences $\Re$. We will denote by $P_i$ and $I_i$ the asymmetric and symmetric factors of $R_i$. The set of all possible strict orderings over the finite set of alternatives $A$ is denoted by $\wp$.

Frequently -when talking about the set of alternatives $[0, M]$- we will assume that the preference relations are single-peaked. A preference relation $R_i$ on $[0, M]$ is *single-peaked* if and only if there exits a unique number $p(R_i) \in [0, M]$ such that $\forall x, y \in [0, M]$, if $y < x \leq p(R_i)$ or $p(R_i) \leq x < y$, then, $xP_iy$. The number $p(R_i)$ will be referred to as the *peak* of agent $i's$ preference relation -exploiting this preferences' evident analogy with a mountain-, since it is, by definition, the most preferred alternative of agent $i$.

When working with $A = [0, M]$ and single-peaked preferences, continuity of preferences is usually assumed. Preferences $R_i \in \Re$ are *continuous* if and only if for every alternative, both the upper and the lower contour sets are closed, i.e., $\forall x \in [0, M] = A$, $\{y \in A \mid yR_ix\}$ and $\{y \in A \mid xR_iy\}$ are closed. This is a natural assumption when dealing with infinite sets and it is sufficient to guarantee that for every closed interval contained in $[0, M]$, there exist a most-preferred alternative. Let $\Re^{SP}$ be set of all continuous and single-peaked preference relations on $A = [0, M]$.

An ordered list of preference relations for all the individuals will be called a *preference profile* and denoted by $\mathbf{R} = (R_i)_{i \in N} = (R_1, ..., R_n)$. We will often use the following notation: given a fixed preference profile $\mathbf{R} = (R_1, ..., R_n)$, $(R', \mathbf{R}_{-i})$ is the profile in which individual $i$ takes preferences $R'$ and any other agent $j \neq i$ remains with the same preferences he had in profile $\mathbf{R}$, i.e., $R_j$. The list $(R', R'', \mathbf{R}_{-i-j})$ is the profile such that the preference relations of agents $i$ and $j$ in profile $\mathbf{R}$, have been replaced by preference relations $R'$ and $R''$ respectively and the other agents' preferences are the same than those they had in profile $\mathbf{R}$. Then, whatever preference relation is placed in the first component of some partitioned profile $(.,.,\mathbf{R}_{-i-j})$ stands for the preference relation of agent $i$ in that profile. Hence, the profile $(R', R'', \mathbf{R}_{-j-i})$ is intended to be the profile $\mathbf{R}$ when agent $j$ has preferences $R'$ and agent $i$ is endowed with preferences $R''$. Moreover, our particular notation admits that some agent's new preference relation is the same preference relation of that of some other agent in the original profile $\mathbf{R}$, in which case we are allowed to refer to that preference relation with its former subscript in order to avoid notation; but notice that the subscript accompanying some individual preference relation in our partitioned notation is *not* related with the agent owning that preference relation in the actual profile, but with the agent that had it in the original -or *reference*- profile. Let us illustrate this important point with an example: The profile $(R_j, R', \mathbf{R}_{-i-j})$ should be read in the following way: "individual $i$ has the same preference relation that individual $j$ had in profile $\mathbf{R}$ $(R_j)$, agent $j$ possesses the preference relation $R'$ and the remaining agents are endowed with the same preference relations they had in the reference profile $\mathbf{R}$".

When preferences are single-peaked, the associated vector of peaks will be: $p(\mathbf{R}) = (p(R_i))_{i \in N} \in [0, M]^n$.

Now, we should model social objectives. A *social choice function* (SCF) is a function which associates a chosen alternative to every economy -or preference profile- and will be denoted by $f : \Re^n \longrightarrow A$.

When we work with the set of alternatives $[0, M]$ and single-peaked preferences, we will be interested in a special kind of SCFs called *voting schemes*. Voting schemes only use information about the agents' peaks, so we can define a voting scheme $\Pi$ as a SCF in which the following holds:

$$\forall \mathbf{R}, \mathbf{R}' \in \Re^n \ s.t.\ p(\mathbf{R}) = p(\mathbf{R}') \Longrightarrow \Pi(\mathbf{R}) = \Pi(\mathbf{R}'). \tag{1}$$

Now we define the properties we shall deal with:

**Definition 1** *For each given $\mathbf{R} \in \Re^n$, $x$ is an **efficient alternative** if $x \in A$ and there is no $x' \in A$ with $x' R_i x$ $\forall i \in N$ and $x' P_i x$ for some $i \in N$. The set of efficient alternatives associated to profile $\mathbf{R}$ will be denoted by $P(\mathbf{R})$*

*A SCF $f$ is efficient if it selects efficient alternatives for each preference profile, i.e., $\forall \mathbf{R} \in \Re^n$, $f(\mathbf{R}) \in P(\mathbf{R})$.*

In the case of $R_i$ single- peaked for all $i \in N$ and $A = [0, M]$, it is easy to prove that $f$ is efficient whenever $\forall \mathbf{R} \in \Re^n$,

$$f(\mathbf{R}) \in P(\mathbf{R}) = [\min\{p(R_i) \mid i \in N\}, \max\{p(R_i) \mid i \in N\}].$$

**Definition 2** *A SCF $f$ is* ***manipulable*** *by agent $i \in N$ at profile $\mathbf{R} \in \Re^n$ via $R_i' \in \Re^n$ if and only if $f(R', \mathbf{R}_{-i}) P_i f(\mathbf{R})$. Whenever a SCF is manipulable by some agent at some profile via a preference relation we say that the SCF is manipulable.*

**Definition 3** *A SCF $f$ is* ***strategy-proof*** *if and only if it is not manipulable.*

This property constitutes a strong incentive compatibility requirement, meaning that agents' lies about their true preferences cannot be in any case profitable -whatever the declared preferences of others may be-. Strategy-proofness may therefore be interpreted as requiring that revealing actual preferences be a dominant strategy for all agents if the SCF is used to choose alternatives based on the agents' reported preferences.

**Definition 4** *A SCF $f$ is* ***anonymous*** *if any permutation of the different values of its arguments yields the same alternative -, i.e., for all one-to-one mappings $\sigma : N \to N$ and all $\mathbf{R} \in \Re^n$, $f(R_1, ..., R_n) = f(R_{\sigma(1)}, ..., R_{\sigma(n)})$.*

This property guarantees that no information about the individuals' names is used in the decision rule.

**Definition 5** *A SCF $f$ satisfies the property of* ***Welfare-domination under preference-replacement*** *(WDUPR)*[2] *if $\forall i \in N$, $\forall \mathbf{R} \in \Re^n$, $\forall R' \in \Re$, either $f(\mathbf{R}) R_j f(R', \mathbf{R}_{-i})$ $\forall j \in N \setminus \{i\}$ or $f(R', \mathbf{R}_{-i}) R_j f(\mathbf{R})$ $\forall j \in N \setminus \{i\}$ .*

The change in the preferences of any individual makes that all the remaining agents *move* in the same welfare direction: either all of them gain or all of them lose -in the weak sense-.

We now introduce two versions of the main condition in this paper. The motivation for both versions is the same and they only differ in what they require when agents are left indifferent when facing somebody's change in preferences. Although both versions are quite similar, the possibilities of finding social choice functions are very different when we require each version to hold, so the apparently slight difference is proved to be crucial in allowing for positive results.

[2] This property has also been called "Replacement monotonicity" and "Replacement domination".

**Definition 6** *A SCF $f$ satisfies the property of* ***strong reciprocity*** *if*
$\forall i \in N, \ \forall \mathbf{R} \in \Re^n, \ \forall R' \in \Re, \ \forall j \in N \backslash \{i\}, \quad f(R', \mathbf{R}_{-i}) R_j f(\mathbf{R}) \Rightarrow f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$.

When agent $i$ changes his preferences - from $R_i$ to $R'$ - and does not affect negatively individual $j$, we require that if $j$ were the one who changed his preferences from the initial profile to the same agent $i's$ new preferences -from $R_j$ to $R'$-, and individual $i$ would remain unchanged -with $R_i$-, $i$ would not lose with $j's$ change either -so $f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$ holds-. Symmetrically, if such a change by individual $i$ makes agent $j$ be worse off -interchanging the roles of $R_i$ and $R'$ above-, the reasoning is the same, but individual $i$ with initial preferences $R'$ should now weakly lose -so that $f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$ holds in this case too-.

**Definition 7** *A SCF $f$ satisfies the property of* ***weak reciprocity*** *if*
$\forall i \in N, \ \forall \mathbf{R} \in \Re^n, \ \forall R' \in \Re, \ \forall j \in N \backslash \{i\}, \ \textit{not}\,[f(\mathbf{R}) P_j f(R', \mathbf{R}_{-i})] \Rightarrow \textit{not}\,[f(\mathbf{R}) P_i f(R', \mathbf{R}_{-j})]$.

In words, if agent $i$ does not make me be (strictly) worse off by changing his preference to $R'$, I should not be able to (strictly) damage him if I would be the agent who changes to $R'$ and $i$ will remain unchanged. Weak reciprocity imposes a ban on perverse asymmetric feelings.

Notice that weak reciprocity implies the following statements: if somebody makes me gain, I can either improve or not affect at all his position (a). If the changing agent is damaging me, again I can either cause him a loss or leaving him unaffected (b). Finally, if I am indifferent with $i'$s change -he has not made me be (strictly) worse off-, the definition applies and I should not damage him: in my (reciprocate) turn, I should be able either to make him gain or break even (c).

$(a)$. $f(R', \mathbf{R}_{-i}) P_j f(\mathbf{R}) \Rightarrow f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$.
$(b)$. $f(\mathbf{R}) P_j f(R', \mathbf{R}_{-i}) \Rightarrow f(\mathbf{R}) R_i f(R', \mathbf{R}_{-j})$.
$(c)$. $f(R', \mathbf{R}_{-i}) I_j f(\mathbf{R}) \Rightarrow f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$ &
$f(R', R_i, \mathbf{R}_{-i-j}) R' f(R', \mathbf{R}_{-i})$.

Weak reciprocity relaxes strong reciprocity in just one sense. We must take a short detour in order to explain the difference. It is not difficult to check that strong reciprocity implies: $\forall i \in N, \ \forall \mathbf{R} \in \Re^n, \ \forall R' \in \Re, \ \forall j \in N \backslash \{i\}$,

$$f(R', \mathbf{R}_{-i}) I_j f(\mathbf{R}) \Rightarrow f(R', \mathbf{R}_{-j}) I_i f(\mathbf{R}) \ \& \ \ f(R', \mathbf{R}_{-i}) I' f(R', R_i, \mathbf{R}_{-i-j})$$

In order to prove this, just note that the indifference on the left side in the former statement is: $f(R', \mathbf{R}_{-i})R_j f(\mathbf{R})$ and $f(\mathbf{R})R_j\ f(R', \mathbf{R}_{-i})$; by applying strong reciprocity to both expression, we get: $\forall i \in N,\ \forall \mathbf{R} \in \Re^n,\ \forall R' \in \Re,\ \forall j \in N \backslash \{i\}$ ,

$f(R', \mathbf{R}_{-i})R_j f(\mathbf{R}) \Rightarrow$

$\Rightarrow f(R', \mathbf{R}_{-j})R_i f(\mathbf{R})$ (1) & $f(R', \mathbf{R}_{-i})R' f(R', R_i, \mathbf{R}_{-i-j})$ (2).

$f(\mathbf{R})R_j\ f(R', \mathbf{R}_{-i}) \Rightarrow$

$\Rightarrow f(R', R_i, \mathbf{R}_{-i-j})R' f(R', \mathbf{R}_{-i})$ (3) & $f(\mathbf{R})R_i f(R_i, R', \mathbf{R}_{-i-j})$ (4).

And since $f(R_i, R', \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-j})$ by definition, (1) and (4) imply $f(R', \mathbf{R}_{-j})I_i f(\mathbf{R})$

and (2) and (3) imply $f(R', \mathbf{R}_{-i})I' f(R', R_i, \mathbf{R}_{-i-j})$, which means that whenever a change from some $i$ leaves an agent $j$ indifferent -for instance, when $i's$ change cannot shift the initial social choice-, the same change from $j$ should not affect $i's$ preferences. This is a stronger requirement than desired, since there might be no reasons for forbidding $i$ to strictly gain, while there may be reasons for $i$ to lose when $j$ change -the reciprocate symmetry might forbid "perverse" hypothetical effects, but there does not seem to be a strong reason to maintain such a strong implication . Weak reciprocity eliminates this requirement by allowing unaffected agents to improve $i's$ position, while not letting him become worse off.[3]

Notice that strong reciprocity always implies weak reciprocity but the converse is not true -(2) and (4) cannot be derived from weak reciprocity-.

**Definition 8** *A SCF $f$ is* ***dictatorial*** *if* $\exists i \in N$ *such that* $\forall R_i \in \Re,\ \forall \mathbf{R}_{-i} \in \Re_{-i},\ \ f(R_i, \mathbf{R}_{-i}) \in \{a \in A \ \ |\ \ aR_i b\ \forall b \in A\}$.

We will need this class of undesirable SCFs in some proofs.

**Definition 9** *A SCF $f$ is* ***constant*** *if* $\exists a \in A$ *such that*
$\forall \mathbf{R} \in \Re^n,\ f(\mathbf{R}) = a.$

**Definition 10** *A SCF $f$ is a* ***Generalized Condorcet winner solution***

[3]There is still other possibility of defining an even weaker concept of reciprocity, consisting on not imposing any constraint at all on the behavior of the rule in indifference situations -and allowing for the "perverse" effect in indifference situations-. The author have explored this possibility but characterisations become much more complicated, although our intuition is that the results obtained with our version would not change very much.

*$(GCWS(n+1))$ if $\exists\alpha = (\alpha_1, \alpha_2, ..., \alpha_{n+1}) \in [0, M]^{n+1}$, called phantom voters or fixed ballots, such that*

$$f(\mathbf{R}) = m\left(p(R_1), p(R_2), ..., p(R_n), \alpha_1, \alpha_2, ..., \alpha_{n+1}\right). \tag{2}$$

*where $m$ stands for the median. Notice that $GCWS(n+1)$ are voting schemes.*

Moulin (1980) showed that when preferences are single-peaked on the interval $[0, M]$, the only anonymous and strategy-proof voting schemes on $[0, M]$ are those belonging to the family of $GCWS(n+1)$. If efficiency is additionally imposed, the resulting class is also the median, but with only $n-1$ phantom voters. We will refer to this family as $GCWS(n-1)$.[4]

**Definition 11** *A SCF $f$ is* ***adjusted constant to*** *$a$ ($a \in [0, M]$) if for all $R \in \Re^n$,*

$$f^a(R) = \begin{cases} a & if\ a \in P(\mathbf{R}) \\ \min\{p(R_i) \mid i \in N\} & if\ a < \min\{p(R_i) \mid i \in N\} \\ \max\{p(R_i) \mid i \in N\} & if\ a > \max\{p(R_i) \mid i \in N\} \end{cases}$$

*Denote by $\Phi$ the family of adjusted-constant SCFs $f$, namely*
*$\Phi = \{f^a \mid a \in [0, M]$ and $f^a$ is adjusted-constant to $a\}$.*

Thomson [13] proved that class $\Phi \subset GCWS(n-1)$ contains the only efficient SCFs such that $WDUPR$ holds when preferences are single-peaked on $[0, M]$. Notice that all the SCFs within class $\Phi$ are anonymous, but not trivial and it is a subclass of the family $GCWS(n-1)$ where we have the $n-1$ phantom voters allocated to the same point.

## 2.3 Results

We will study the behavior of the reciprocity property under two different domain assumptions. First, we characterize the anonymous and strong reciprocate SCFs in the *unrestricted domain* of every preference relation when the set of alternatives is finite and we will obtain a result that establishes a close relationship between strategy-proofness and both reciprocity and anonymity. We will benefit from this property to prove the characterization result by

[4] The median for the case of $n+1$ phantom voters is defined as:
$m\left(p(R_1), p(R_2), ..., p(R_n), \alpha_1, \alpha_2, ..., \alpha_{n+1}\right) \Leftrightarrow$
$\#\{i \mid p(R_i) \leq m\} + \#\{i \mid \alpha_i \leq m\} \geq n$ and
$\#\{i \mid p(R_i) \geq m\} + \#\{i \mid \alpha_i \geq m\} \geq n$.

means of the well-known Gibbard-Satterthwaite Theorem. The negative result shows the impossibility of finding strong reciprocate and anonymous SCFs in this domain.

Secondly, we investigate the existence of anonymous, and strong/weak reciprocate SCFs in contexts where preferences are restricted to satisfy single-peakedness on the closed interval of the real line. We can use some theorems related with strategy-proof SCFs with single-peaked preferences: Moulin's [6] characterization of strategy-proof voting schemes and the extensions of this result to general SCFs: Barberà & Jackson [3], Barberà, Sonnenschein & Zhou [2], that will allow us to use the relation between strategy-proofness and reciprocity. The characterization theorems in this case results in an impossibility for anonymous, efficient and strong reciprocate SCFs and the $GCWS$ voting schemes are shown to be the only anonymous, efficient and weak reciprocate SCFs. Before establishing the main results in this section, we need to prove two useful lemmata. *Lemma 12*, is interesting on its own, since it shows that for any domain of preferences, weak reciprocity and anonymity together imply that if someone's change makes me strictly gain, I cannot make him be strictly better in my turn, but if someone strictly worsens my position, I can be sure that I would make him lose if I were the one who changed. The second lemma, *Lemma 13*, has no an easy interpretation and is intended to simplify some proofs.

**Lemma 12** *Assume $f$ is a weak reciprocate and anonymous SCF. Then,* $\forall i \in N,\ \forall \mathbf{R} \in \Re^n,\ \forall R' \in \Re,\ \forall j \in N\backslash\{i\}$,

$$f(R', \mathbf{R}_{-i}) P_j f(\mathbf{R}) \Rightarrow f(R', \mathbf{R}_{-j}) I_i f(\mathbf{R})\ \&\ \ f(R', \mathbf{R}_{-i}) P' f(R', R_i, \mathbf{R}_{-i-j}).$$

**Proof.** Take any $i \in N$, $\mathbf{R} \in \Re^n$, $R' \in \Re$ and $j \in N\backslash\{i\}$, and suppose that $f(R', \mathbf{R}_{-i}) P_j f(\mathbf{R})$ (1). Since $f$ is weak reciprocate, consider first the profile $\mathbf{R} = (R_i, R_j, \mathbf{R}_{-i-j})$ and suppose that individual $i$ changes his preferences from $R_i$ to $R'$, reaching the profile $(R', R_j, \mathbf{R}_{-i-j})$. By (1), individual $j$ strictly gains with $i's$ change, so by weak reciprocity, if it were individual $j$ who was suffering the same change instead of $i$, the latter individual would not lose, so it holds that $f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$ (2). Now, consider that the initial profile is $(R', \mathbf{R}_{-i})$ and agent $i$ changes his preferences to $R_i$ - the converse of the former shift -. By condition (1), individual $j$ worsens his position, so by weak reciprocity, if $j$ were the agent who changed to preference $R_i$ - abstracting from the subscript - and $i$ would remain unchanged, he would be weakly worsened likewise, so it also holds that $f(R', \mathbf{R}_{-i})R'f(R', R_i, \mathbf{R}_{-i-j})$ (3). We know till now that (2) and (3) hold, but there may be two possibilities in each of those conditions: Each can be satisfied with strict preference or with indifference. We will denote every possibility as: $(2P), (2I), (3P)$ and $(3I)$, i.e., $f(R', \mathbf{R}_{-j})P_i f(\mathbf{R})$ $(2P)$ $\quad f(R', \mathbf{R}_{-i})P'f(R', R_i, \mathbf{R}_{-i-j})$ $(3P)$

$f(R', \mathbf{R}_{-j})I_i f(\mathbf{R})$ $(2I)$ $\quad f(R', \mathbf{R}_{-i})I'f(R', R_i, \mathbf{R}_{-i-j})$ $(3I)$

Now, we will check all the combined possibilities:

**1**- $(2P)$ and $(3P)$ :

Let us consider $(2P)$ and focus on the profile $(R', \mathbf{R}_{-j}) = (R_i, R', \mathbf{R}_{-i-j})$. Now, imagine that individual $j$ with preferences $R'$ changes to preferences $R_j$, so that the final profile will be $(R_i, R_j, \mathbf{R}_{-i-j}) = \mathbf{R}$. Since we are assuming that $(2P)$ holds, agent $i$ would be worse off, so by weak reciprocity, if $i$ would had changed from preferences $R_i$ in profile $(R_i, R', \mathbf{R}_{-i-j})$ to $R_j$, $i$ remaining unchanged, he could not improve agent $i's$ situation. Hence,

$f(R_i, R', \mathbf{R}_{-i-j})R'f(R_j, R', \mathbf{R}_{-i-j})$ (4). Now, by anonymity, agents' names do not matter, so, we have: $f(R_i, R', \mathbf{R}_{-i-j}) = f(R', R_i, \mathbf{R}_{-i-j})$ and

$f(R_j, R', \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-i})$, and we can write (4) as:

$f(R', R_i, \mathbf{R}_{-i-j})R'f(R', \mathbf{R}_{-i})$. Notice that this last expression directly contradicts $(3P)$, so the present possibility cannot appear.

**2**- $(2P)$ and $(3I)$ :

Let us focus on $(3I)$ and profile $(R', \mathbf{R}_{-i}) = (R', R_j, \mathbf{R}_{-i-j})$ : Suppose that individual $j$ changes his preferences to $R_i$ -agent $i's$ preferences in profile $\mathbf{R}$- so that the final situation is $(R', R_i, \mathbf{R}_{-i-j})$. $(3I)$ implies, then, that individual $i$ with preferences $R'$ is indifferent about the shift, so by weak reciprocity, if he were the one who changed to preferences $R_i$, he could not have worsened individual $j's$ position with preferences $R_j$, which can be written as: $f(R_i, R_j, \mathbf{R}_{-i-j}) = f(\mathbf{R})R_j f(R', \mathbf{R}_{-i}) = f(R', R_j, \mathbf{R}_{-i-j})$ (5). Notice that this statement contradicts directly the assumption (1), so this case is impossible.

**3**- $(2I)$ and $(3I)$ :

This case cannot occur either, since it is identical to case 2 in the sense that only $(3I)$, when present, causes the contradiction with (1) whether the case is $(2P)$ or $(2I)$.

**4**- $(2I)$ and $(3P)$ :

This turns out to be the only possibility allowed by both weak reciprocity and anonymity, so the lemma is proved. ■

**Lemma 13** *Assume $f$ is a strong reciprocate and anonymous SCF. Then,*
$\forall i \in N$, $\forall \mathbf{R} \in \Re^n$, $\forall R' \in \Re$ *such that* $f(R', \mathbf{R}_{-i}) \neq f(\mathbf{R}) \Rightarrow$
$f(R', \mathbf{R}_{-i}) I' f(R', \mathbf{R}_{-j}) \ \ \forall j \in N \backslash \{i\}$.

**Proof.** Suppose any $i \in N$, $\mathbf{R} \in \Re^n$, $R' \in \Re$ such that $f(R', \mathbf{R}_{-i}) \neq f(\mathbf{R})$. Then, let us take some individual other than the one who shifted the decision ($i$), for example, agent $j$ and find out in what direction he was affected by $i$th's shift from $R_i$ to $R'$. there are two possibilities: either $f(R', \mathbf{R}_{-i}) R_j f(\mathbf{R})$ or $f(\mathbf{R}) R_j f(R', \mathbf{R}_{-i})$. We will distinguish both cases:

*Case 1*: $f(R', \mathbf{R}_{-i}) R_j f(\mathbf{R})$ (1). Consider now the change of agent $i$ from preferences $R_i$ to $R'$ : by assumption, $j$ does not loose. We can use strong reciprocity with respect to agent $i$ and obtain: $f(R', \mathbf{R}_{-j}) R_i f(\mathbf{R})$ (2). Consider now the profile $(R', \mathbf{R}_{-j})$ and suppose that agent $j$ with preferences $R'$ changes to his original one $(R_j)$. We come back to the profile $\mathbf{R}$. By expression (2), agent $i$ weakly looses, and by strong reciprocity $f(R_i, R', \mathbf{R}_{-i-j}) R' f(R_j R', \mathbf{R}_{-i-j})$ (3). But, by anonymity, any permutation of the arguments of the SCF cannot modify its value, and the following will hold: $f(R', R_j, \mathbf{R}_{-i-j}) = f(R_j, R', \mathbf{R}_{-i-j})$. We can, then, rewrite expression (3) in this way:

$f(R_i, R', \mathbf{R}_{-i-j}) R' f(R', R_j, \mathbf{R}_{-i-j})$ $(3')$.

Let us focus now on profile $(R', R_j, \mathbf{R}_{-i-j})$ and imagine that agent $i$ changes preferences $R'$ to $R_i$ -the converse of the initial change-. By hypothesis (1), $j$ should be in a worse position, so by strong reciprocity, individual $i$ should move in the same direction if $j$ were the one who changed. In other words, the following holds true: $f(R', R_j, \mathbf{R}_{-i-j}) R' f(R', R_i, \mathbf{R}_{-i-j})$ (4). Again by anonymity, permuting preferences of agents yields the same social choice, and this holds: $f(R', R_i, \mathbf{R}_{-i-j}) = f(R_i, R', \mathbf{R}_{-i-j})$. Expression (4) can be expressed this way: $f(R', R_j, \mathbf{R}_{-i-j}) R' f(R_i, R', \mathbf{R}_{-i-j})$ $(4')$. Statements $(3')$ and $(4')$ are obtained from the assumptions, and both should be simultaneously true, so in this case we conclude: $f(R', R_j, \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-i}) R' f(R', \mathbf{R}_{-j}) = f(R_i, R', \mathbf{R}_{-i-j})$.

*Case 2*: $f(\mathbf{R}) R_j f(R', \mathbb{R}_{-i})$. We can follow the same steps as in case 1. The only difference is that every preference relation is inverted, and we obviously reach the same conclusion as in case 1.

**Corollary 14** *Let $A$ be a finite set of alternatives and $\Re = \wp$. If SCF $f$ is strong reciprocate and anonymous, then, $f$ is strategy-proof.*

**Proof.** We prove it by contradiction: we suppose that $f$ is not strategy-proof but it is both strong reciprocate and anonymous, and we will find a contradiction. If $f$ is not strategy-proof, then, there exist: $\exists i \in N$, $\exists \mathbf{R} \in \Re^n$, $\exists R' \in \Re$, such that $f(R', \mathbf{R}_{-i}) P_i f(\mathbf{R})$. This obviously implies that $f(R', \mathbf{R}_{-i}) \neq f(\mathbf{R})$, so we can directly apply *Lemma 13* and obtain

$f(R', \mathbf{R}_{-i}) I' f(R', \mathbf{R}_{-j}) \quad \forall j \in N \backslash \{i\}$ (1). Since we are working with strict orderings, it implies that

$f(R', \mathbf{R}_{-i}) = f(R', \mathbf{R}_{-j}) \ \ \forall j \in N \backslash \{i\}$. Consider now the change of individual $j$ with preferences $R_i$ in profile $(R', \mathbf{R}_{-j}) = (R', R_i, \mathbf{R}_{-i-j})$ to his original preferences $R_j$ , reaching profile $(R', R_j, \mathbf{R}_{-i-j})$. From (1), individual $i$ with preferences $R'$ remains indifferent with the change, so by strong reciprocity and anonymity, we get $f(R', \mathbf{R}_{-i}) = f(\mathbf{R})$, contradicting our initial assumption. ■

**Corollary 15** *Let $A$ be a finite set of alternatives and $\Re = \wp$. There do not exist anonymous and strong reciprocate SCFs. such that $\#(range(f)) \geq 3$.*

The proof is obvious by using *Corollary 13* and the Gibbard-Satterthwaite Theorem -Gibbard [5], Satterthwaite [9]-.

The former negative result leads us either to consider more restricted domains of preferences or to focus on weak reciprocity. If we relax the reciprocity condition to its weaker version, we can see that there exist efficient, anonymous and weak reciprocate SCFs, even for quite rich domains, like the one of *strict orderings* over alternatives. Let us consider $n = 3$, $A = \{a, b, c\}$ and the following class of SCFs.: $\forall \mathbf{R}^3 \in \Re^3 = \wp^3$,

$$f^a(R) = \begin{cases} a & if \quad a \in P(\mathbf{R}) \\ b & if \quad a \notin P(\mathbf{R}) \ \& \ D(b, c, \mathbf{R}) > D(c, b, \mathbf{R}) \\ c & if \quad a \notin P(\mathbf{R}) \ \& \ D(b, c, \mathbf{R}) < D(c, b, \mathbf{R}) \end{cases}$$

where $D(x, y, \mathbf{R}) = \#\{i \in N \ s.t. \ xR_i y\} \ \ \forall x, y \in A, \ \ \forall \mathbf{R} \in \Re^3$. It is not difficult to prove that this SCF and its analogous are efficient, anonymous, weak reciprocate and satisfy $WDUPR$. Unfortunately, they weight excessively an arbitrary status quo, they are not strategy-proof and it is not clear how they can be generalized to more than three alternatives or to domains admitting indifference sets. Therefore, We will now consider other domains.

In order to compare both versions of reciprocity with $WDUPR$, we focus on the restricted domain of continuous, single-peaked preference relations on $A = [0, M]$ . From now on, we should distinguish between both kinds of reciprocity, which will be separately explored. We start with our main results concerning strong reciprocity.

**Theorem 16** *Let* $\Re = \Re^{SP}$ *and* $n = 2$. *Then, there do not exist efficient, strong reciprocate and anonymous SCFs.*

**Proof.** Let us consider any profile $R = (R_1, R_2)$ such that $P(R_1) = 0$ and $P(R_2) = M$. Suppose w.l.g. that $f(R_1, R_2) \in [0, M)$ - otherwise, just permute the names of the agents and the reasoning will be analogous -. Now, consider any profile $\widehat{R}_1$ such that $P(\widehat{R}_1) \in (f(R_1, R_2), M]$ and such that $\forall x \geq P(\widehat{R}_1),\ x\widehat{R}_1 y\ \forall y < P(\widehat{R}_1)$. - Notice that there always exist admissible single-peaked preferences for which that condition holds -. Now, suppose that individual 1 in profile $R$ changes his initial preferences $R_1$ to preferences $\widehat{R}_1$, such that the new profile will be $(\widehat{R}_1, R_2)$. Since there are just two agents, efficiency requires that $f(\widehat{R}_1, R_2) \in \left[P(\widehat{R}_1), P(R_2)\right]$ , so $f(\widehat{R}_1, R_2) > f(R_1, R_2)$ and single-peaked preferences makes agent 2 in profile $R$ with preferences $R_2$ be strictly better off with $i's$ change, since $f(\widehat{R}_1, R_2) P_2\ f(R_1, R_2)$, so by strong reciprocity, if agent 2 were the one who changed to preferences $\widehat{R}_1 = \widehat{R}_2$ while agent 1 would remain unchanged with $R_1$, $f(R_1, \widehat{R}_2) R_1$ $f(R_1, R_2)$. Hence, since $P(R_1) = 0$, it must be that $f(R_1, \widehat{R}_2) \leq f(R_1, R_2)$.

Now, let us consider profile $(R_1, \widehat{R}_2)$ and suppose that agent 2 with preferences $\widehat{R}_2$ changes to new preferences $\widehat{\widehat{R}}_2 = R_2$, the new profile being $(R_1, R_2)$. since we know from above that $f(R_1, \widehat{R}_2) \leq f(R_1, R_2) = f(R_1, \widehat{\widehat{R}}_2)$, individual 1 with preferences $R_1$ can either be indifferent with the change whenever $f(R_1, \widehat{R}_2) = f(R_1, R_2)$ or strictly loose if the case is that of $f(R_1, \widehat{R}_2) < f(R_1, \widehat{\widehat{R}}_2) = f(R_1, R_2)$. Suppose first that $f(R_1, \widehat{R}_2) < f(R_1, R_2)$ : this implies that 1 loses with the change, and strong reciprocity requires that, if he were the one who changed from $R_1$ to $\widehat{\widehat{R}}_1 = \widehat{\widehat{R}}_2 = R_2$, agent 2 with initial preferences $\widehat{R}_2$ in profile $(R_1, \widehat{R}_2)$ could never gain with the change. Therefore,

$$f(R_1, \widehat{R}_2) \widehat{R}_2 f(\widehat{\widehat{R}}_1, \widehat{R}_2) = f(R_2, \widehat{R}_2). \tag{3}$$

Since, by anonymity, $f(R_2, \widehat{R}_2) = f(\widehat{R}_2, R_2) = f(\widehat{R}_1, R_2)$, the last expression can be rewritten as $f(\widehat{R}_1, R_2)\widehat{R}_2 f(\widehat{R}_1, R_2) \in \left[P(\widehat{R}_1), P(R_2)\right]$. But by definition of $\widehat{R}_1$,

$\forall x \in \left[P(\widehat{R}_1), P(R_2)\right]$, $x\widehat{P}_1 y$ $\forall y \in \left[0, P(\widehat{R}_1)\right)$, so since $\widehat{R}_2 = \widehat{R}_1$ and $f(R_1, \widehat{R}_2) \leq f(R_1, R_2) < P(\widehat{R}_1) \leq P(R_2) = M$, $f(\widehat{R}_1, R_2)\widehat{P}_1 f(\widehat{R}_1, R_2)$, a contradiction.

It remains to check the case in which $f(R_1, \widehat{R}_2) = f(R_1, R_2)$ and agent 1 is indifferent with $2's$ change from $\widehat{R}_2$ to $\widehat{\widehat{R}}_2 = R_2$. By strong reciprocity, 1 should leave agent 2 indifferent if he were the one who changed preferences, so again by anonymity, it should hold that $f(\widehat{R}_1, R_2) = f(R_1, \widehat{R}_2) < P(\widehat{R}_1) \leq P(R_2)$, contradicting efficiency of $f$ at $f(\widehat{\widehat{R}}_1, \widehat{R}_2) = f(R_2, \widehat{R}_2) = f(\widehat{R}_1, R_2)$. ■

**Theorem 17** *Let $\Re = \Re^{SP}$ and $n \geq 3$. Then, the only strong reciprocate and anonymous social SCFs are constant.*

**Proof.** We have to prove both implications:

*Step 1*:( $\Rightarrow$) $\Re$ single-peaked, $f$ is a strong reciprocate and anonymous SCF with $\#N \geq 3 \Rightarrow f$ is constant.

We will first demonstrate that under the single-peakedness assumption and $\#N \geq 3$, every strong reciprocate and anonymous SCF has to be strategy- proof. It will be proved by contradiction: we first suppose that $f$ is anonymous, manipulable and strong reciprocate and we will find a contradiction.

Let us consider only three ordered individuals to simplify the notation of the proof, and let us call them 1, 2 and 3. The profile that is supposed to be manipulated will be now the following: $(R_1, R_2, R_3)$ and let agent 1 -without loss of generality- be the manipulator, changing to preferences $R'$.

We can now apply *Lemma 13* for $j = 2, 3$ with the above change, so the following statements are true: $f(R_1, R', R_3)I'f(R', R_2, R_3)$ (1) and

$f(R_1, R_2, R')I'f(R', R_2, R_3)$ (1′).

Consider now the change consisting of changing agent 1's preferences from $R_1$ to $R_3$ in the profile $(R_1, R_2, R')$, the final preference profile being: $(R_3, R_2, R')$. By anonymity, this profile has the same value that: $(R', R_2, R_3)$. By expression (1′) both values are considered indifferent with preferences $R'$. Hence, by strong reciprocity in the two directions with respect to agent 1, the following holds true: $f(R_1, R_2, R_3)I_1 f(R_1, R_2, R')$ (2).

Now, let us remember that, by the manipulability hypothesis at the original profile, it is true that: $f(R', R_2, R_3)P_1f(R_1, R_2, R_3)$ (3) -in the notation of *Theorem 16*-. Notice that we assumed that we have at least three individuals, so (1) and (1′) can be written:

$$f(R_1, R_2, R')I'f(R_1, R', R_3)I'f(R', R_2, R_3).$$

But notice that single- peaked preferences only allow for at most two distinct indifference points, so only two possibilities can occur:

**1**- $f(R', R_2, R_3) = f(R_1, R', R_3)$, -or $f(R', R_2, R_3)P_1f(R_1, R_2, R')$- in which case, using the analogous to expression (2) corresponding to the change from $R_1$ to $R_2$ in the profile $(R_1, R', R_3)$, the final preference profile being: $(R_2, R', R_3) : f(R_1, R_2, R_3)I_1f(R_1, R', R_3)$ (4), and the following expression will hold:

$f(R_1, R_2, R_3)I_1f(R_1, R', R_3) = f(R', R_2, R_3)$. This contradicts directly the manipulability hypothesis -expression (1)-.

**2**- either $f(R_1, R', R_3) < f(R', R_2, R_3) < f(R_1, R_2, R_3)$ or

$f(R_1, R_2, R_3) < f(R', R_2, R_3) < f(R_1, R', R_3)$ and always: $f(R_1, R_2, R') = f(R_1, R', R_3)$. Because if $f(R', R_2, R_3) = f(R_1, R_2, R_3)$, there is a contradiction with the manipulability of the original profile, and it is the only possibility for (4) to hold true due to the single-peakedness of preferences. Notice that in this case we can consider profile: $(R_1, R', R_3)$ and suppose that agent 1 changes his preferences from $R_1$ to $R_2$, reaching the profile of preferences $(R_2, R', R_3)$. Let us examine the effect of the change on agent 2 -with preferences $R'$-. As expression (1) holds and, by anonymity, $f(R_2, R', R_3) = f(R', R_2, R_3)$ - permuting agents' 1 and 2 preferences -, by reciprocity with respect to 2 the following relation should be true:

$f(R_1, R', R_3)I_1f(R_2, R_1, R_3) = f(R_1, R_2, R_3)$ - by anonymity -.

Consider now the profile $(R_1, R_2, R_3)$ and suppose that agent 3 with preferences $R_3$ moves to preferences $R'$. The final profile will be $(R_1, R_2, R')$, and by (2), the effect on agent 1 will be: $f(R_1, R_2, R_3)I_1f(R_1, R_2, R')$. Using strong reciprocity and anonymity we have:

$f(R', R_2, R_3) = f(R', R_2, R_3)I_3\ f(R_1, R_2, R_3)$ (5). But $f(R_1, R', R_3)$ is strictly on the right or strictly on the left of $f(R', R_2, R_3)$ and $f(R_1, R_2, R_3)$ and the peak of $R_3$ is such that: $p(R_3) \in [f(R_1, R_2, R_3),\ f(R', R_2, R_3)]$, so single-peakedness will imply:

$f(R', R_2, R_3)P_3\ f(R_1, R', R_3) = f(R_1, R_2, R')$ (6) and

$f(R_1, R_2, R_3)P_3\ f(R_1, R', R_3) = f(R_1, R_2, R')$ (7).

We can construct the symmetric change ($R_2$ moves to preferences $R'$) to check the relation:

$f(R', R_2, R_3)P_2\ f(R_1, R_2, R') = f(R_1, R', R_3)$ $(6')$ and

$f(R_1, R_2, R_3)P_2\ f(R_1, R_2, R') = f(R_1, R', R_3)$ $(7')$.

Let us remember that we are in the only case allowed by the single-peakedness assumption in which: $f(R_1, R_2, R') = f(R_1, R', R_3)$. As both profiles achieve the same social choice, everybody will feel indifferent between them, and in particular, agents with preferences $I' : f(R_1, R_2, R')I'$ $f(R_1, R', R_3)$. This can be written, by anonymity, in this way:

$f(R_1, R_2, R')I'\ f(R_1, R_3, R')$. Let us consider the first profile in the relation and suppose that agent 2 changes his preferences from $R_2$ to $R_3$, obtaining $f(R_1, R_3, R')$. By strong reciprocity with respect to agent 3, the following will be true:

$f(R_1, R_2, R_3)I_2\ f(R_1, R', R_3) = f(R_1, R_3, R')$. But recalling expression $(6')$ and relation (5) for the symmetric case when agent 2 changes from preferences $R_2$ to $R' : f(R', R_2, R_3)I_2\ f(R_1, R_2, R_3)$ (6). From $(6')$ and $(7')$, it should be true that:

$f(R_1, R', R_3)I_2\ f(R_1, R_2, R_3)$. But we have seen that the following is true:

$f(R_1, R_2, R_3)P_2\ f(R_1, R', R_3)$, and this is the contradiction we were looking for.

We have proved till now that under our assumptions, every strong reciprocate and anonymous SCF has to be strategy-proof. Using now Moulin's [6] characterization of anonymous, strategy-proof voting schemes and the results related for SCFs in the right direction, we obtain that such SCFs. should belong to the class of voting schemes defined by Moulin as *Generalized Condorcet winner solutions* $(n+1)$. Now, it suffices to prove that the only voting schemes belonging to the class of $GCWS(n+1)$ that are strong reciprocate are those that allocate all the phantom voters to the same point, i.e.,

$$\{\Pi : \Re^n \to A \mid \Pi(\mathbf{R}) = m\left(p(R_1), p(R_2), ..., p(R_n), a, a, ..., a_{n+1}\right)\}$$

Suppose that we face a voting scheme from the $GCWS(n+1)$ family such that there exist at least two phantom voters allocated in different points in the interval: $\exists \alpha_h, \alpha_l$ with $\alpha_h \neq \alpha_l$, $\alpha_h < \alpha_l$. Take, then any piece of the interval $[\alpha_h, \alpha_l]$ with no phantom voters in it and fix any profile with all the people's peaks inside that interval. It is not difficult to check that the social choice will coincide some of the agents' peaks, say individual $i$ $(p(R_i) = m(x, \alpha))$. Consider that the agent which peak is closer to that of $i$

-let's call him $j$- changes his preferences to any other with peak in the open interval between the initial peaks of $i$ and $j$. It is straightforward that the median cannot change, so everybody feels indifferent with both profiles. By strong reciprocity, if agent $i$ would change to $j's$ new peak and $j$ would be the initial 1, $i$ should be indifferent with both profiles. But this is impossible, since the new social choices changes and cannot jump over anybody's peaks, so $j$ would strictly gain and the voting scheme is not strong reciprocate.

The only voting schemes allowed are, then those with all the $n+1$ phantom voters located at the same point; but this is another expression for the constant function.

*Step 2.* ($\Leftarrow$) Any constant SCF is strong reciprocate. This part is obvious and follows directly from the definition of strong reciprocity. ■

This result turns out to be even worse than expected, since constant SCFs are far more undesirable that Thomson's family $\Phi$, which are at least efficient, so we can fear about the possibilities of introspective solidarity against $WDUPR$. Notice, however, that the apparently narrow behavior of strong reciprocity is extremely sensitive to the unnecessary and strong requirement that we have already seen hiding behind the definition related to the responsiveness of strong reciprocity when facing indifference situations. In this line, we hope that weak reciprocity will yield better results than its stronger version. The problem is that the last proof cannot be applied to the weak reciprocity case because we have used the indifference features that are not shared by weak reciprocity. The analysis can however be simplified when we impose the additional property of efficiency. In exchange, we can forget about the minimal 3-agents size of society of the former result.

**Theorem 18** *Let $\Re = \Re^{SP}$. The only weak reciprocate, anonymous and efficient SCFs are those belonging to the class $GCWS(n-1)$.*

**Proof.** As this is a characterization theorem, we must prove both directions:

$\Rightarrow$) First, we will prove that if $f$ is a weak reciprocate, anonymous and efficient SCF, it has to be strategy-proof. We will proceed by contradiction. Suppose that $f$ is not strategy-proof, but it is weak reciprocate, anonymous and efficient and we will find a contradiction.

If $f$ is not strategy-proof, we know that there exist: $\exists l \in N$, $\exists \mathbf{R} \in \Re^n$, $\exists R' \in \Re$, such that $f(R', \mathbf{R}_{-l}) P_l f(\mathbf{R})$. Since we suppose $f$ to be anonymous, we can rename the individuals and the new SCF will be invariant, so consider the following permutation of agents such that all are reordered according to the following rule: $\forall j, h \in N$,

*if* $f(R', \mathbf{R}_{-i}) < f(\mathbf{R})$, $\sigma(j) < \sigma(h) \Leftrightarrow f(R', \mathbf{R}_{-i}) - p(R_j) > f(R', \mathbf{R}_{-i}) - p(R_h)$.

*if* $f(R', \mathbf{R}_{-i}) > f(\mathbf{R})$, $\sigma(j) > \sigma(h) \Leftrightarrow f(R', \mathbf{R}_{-i}) - p(R_j) < f(R', \mathbf{R}_{-i}) - p(R_h)$.

We can always construct the above permutation, which simply consists in ordering the individuals in direct relation with the distance from his peak to the extreme defined by the direction of the shift in the value of $f$ due to the considered manipulation. Hence, call $i = \sigma(l)$ -the new name of the agent manipulating the rule-, and suppose without loss of generality that $f(R', \mathbf{R}_{-i}) < f(\mathbf{R})$ -all the argument can be easily replicated to the other case-. Now, by efficiency, somebody in the manipulable profile $\mathbf{R}$ should loose with the shift, and moreover, $\exists h > i$ such that $p(R_h) \geq f(\mathbf{R})$, since if not, $f(\mathbf{R})$ would not be an efficient alternative for $\mathbf{R}$ -everybody's peaks would be strictly on the left of $f(\mathbf{R})$-, so take the agent with the highest peak in profile $\mathbf{R}$ -if there are more than one, take any of them- and let us call him $j$, so it holds that $p(R_j) = \max_{h \in N} p(R_h)$. It holds for this individual that $f(\mathbf{R}) P_j f(R', \mathbf{R}_{-i})$ (1). Since $f$ is weak reciprocate and anonymous, applying *Lemma 12* to (1),-while inverting roles of $R_i$ and $R'$- we know that the following statements are true: $f(R', \mathbf{R}_{-i}) I' f(R', R_i, \mathbf{R}_{-i-j})$ (2) and $f(\mathbf{R}) P_i f(R_i, R', \mathbf{R}_{-i-j})$ (3). By linking the manipulability hypothesis with (3), we get: $f(R', \mathbf{R}_{-i}) P_i f(\mathbf{R}) P_i f(R_i, R', \mathbf{R}_{-i-j})$ $(3')$, so the former three profiles yield different outcomes and, by the single-peakedness assumption, only two possibilities can occur:

*(i).* $f(R_i, R', \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-j}) < f(R', \mathbf{R}_{-i}) < f(\mathbf{R})$.

Notice that in this case, it must be by (2) and $(3')$ that

$p(R') \in (f(R', \mathbf{R}_{-j}),\ f(R', \mathbf{R}_{-i}))$, $p(R_i) \in (f(R', \mathbf{R}_{-j}),\ f(\mathbf{R}))$and

$p(R_j) \geq f(\mathbf{R})$. By efficiency of $f(R', \mathbf{R}_{-j})$, there exists some other individual $h$ with preferences in $\mathbf{R}$ such that $p(R_h) \leq f(R', \mathbf{R}_{-j})$, so by single-peakedness, $f(R', \mathbf{R}_{-j}) P_h f(R', \mathbf{R}_{-i})$ $(1')$. Now, consider the change of individual $i$ from preferences $R_i$ in profile $(R_i, R', \mathbf{R}_{-i-j}) = (R', \mathbf{R}_{-j})$ to preferences $R_j$, such that the final profile will be $(R_j, R', \mathbf{R}_{-i-j})$. By anonymity, $f(R_j, R', \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-i})$, and individual $i$ gains with the change by $(3')$, the first profile being manipulable by $i$. Let us check the effect of the shift on individual $h$ : by ($1'$), agent $h$ strictly loses and, by reciprocity with respect to $h$, it holds that $f(R_i, R', R_j, \mathbf{R}_{-i-j-h}) = f(R', \mathbf{R}_{-h}) I_j f(R', \mathbf{R}_{-i})$, ($2'$), and $f(R_i, R', \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-j}) P_i f(R', \mathbf{R}_{-h})$. Since $f(R', \mathbf{R}_{-h}) \neq f(R', \mathbf{R}_{-i})$ by the assumption of manipulability and single-peakedness, it is true that $f(R', \mathbf{R}_{-h}) > f(R', \mathbf{R}_{-i})$, and by ($2'$) and the above restrictions on the peaks, we know that $p(R_j) \in (f(R', \mathbf{R}_{-i}),\ f(R', \mathbf{R}_{-h}))$ and hence,

$p(R_i), p(R_j), p(R') <\ f(R', \mathbf{R}_{-h})$. But, notice that by construction,

$p(R_j) = \max_{g \in N} p(R_g)$, so it must be that

$f(R', \mathbf{R}_{-h}) > \max \left\{ \max_{g \in N} p(R_g),\ p(R') \right\}$, so $f(R', \mathbf{R}_{-h})$ cannot be an efficient alternative in profile $(R', \mathbf{R}_{-h})$, a contradiction.

($ii$). $f(R', \mathbf{R}_{-i}) < f(\mathbf{R}) < f(R_i, R', \mathbf{R}_{-i-j}) = f(R', \mathbf{R}_{-j})$, so if both extreme profiles are considered indifferent by preferences $R'$ by expression (2), since single-peaked indifferent sets have at most two points, it must be that $f(\mathbf{R}) P' f(R', \mathbf{R}_{-j})$ (4) and $f(\mathbf{R}) P' f(R', \mathbf{R}_{-i})$ ($4'$). Now, let us consider the profile $(R_i, R', \mathbf{R}_{-i-j})$; Notice that (4) and (3) respectively imply: $p(R') < f(R', \mathbf{R}_{-j})$ and $p(R_i) < f(R', \mathbf{R}_{-j})$, so by efficiency of $f(R', \mathbf{R}_{-j})$, $\exists j' \in N,\ j' \neq j$, such that $p(R_{j'}) \geq f(R', \mathbf{R}_{-j})$. Now, let us consider that individual $j$ with preferences $R'$ in profile $(R_i, R', \mathbf{R}_{-i-j})$ changes to preferences $R_j$ -his initial ones-, reaching the profile $R$. By expression (4), $j$ with preferences $R'$ strictly gains by declaring $R_j$, so we have found another manipulable profile. Moreover, agent $j'$ strictly looses with the change, so we also know that: $f(R_i, R', \mathbf{R}_{-i-j}) P_{j'} f(\mathbf{R})$. Now, we are in the conditions of applying *Lemma 12* and repeating all the former steps again, where only the case ($ii$) is to be considered, but now the role of preferences $R_i$ is performed by $R'$, the role of $R'$ is carried out by $R_j$ and the one of $R_j$ is for $R_{j'}$, so we can always construct a sequence of profiles of the form:

$$\begin{aligned}
&\mathbf{R}^{(1)} = (R_1, R', R_3, R_4, R_5, ..., R_n)\\
&\mathbf{R}^{(2)} = (R_1, R_2, R', R_4, R_5, ..., R_n)\\
&\mathbf{R}^{(3)} = (R_1, R_2, R_3, R', R_5, ..., R_n)\\
&\mathbf{R}^{(4)} = (R_1, R_2, R_3, R_4, R', ..., R_n)\\
&.......................................\\
&\mathbf{R}^{(n-1)} = (R_1, R_2, R_3, R_4, R_5, ..., R')
\end{aligned}$$

in which some agent can manipulate the rule by changing preferences to another initially present in profile $R$ and such that:

$f(\mathbf{R}^{(1)}) < f(\mathbf{R}^{(2)}) < f(\mathbf{R}^{(3)}) < f(\mathbf{R}^{(4)}) < ... < f(\mathbf{R}^{(n-1)})$ and
$\forall h \in \{1, ..., n-1\}, \ \forall l \in \{1, ..., h\}, \ \ f(\mathbf{R}^{(h)}) > p(R') \geq \max_l p(R_l)$.

Therefore, profile $\mathbf{R}^{(n-1)}$ cannot be efficient, since there are no more individuals with preferences in $\mathbf{R}$ with peaks on the right of $f(\mathbf{R}^{(n-1)})$ and every peak is strictly on the left of $f(\mathbf{R}^{(n-1)})$. This is a contradiction and $f$ has to be strategy-proof. Now, we can apply Barberà & Jackson [3] result: the only strategy-proof SCFs must be voting schemes, and Moulin's [6] result, which states that every anonymous, efficient and strategy-proof voting scheme should belong to the family of $GCWS(n-1)$.

$\Leftarrow$) The implication: $\Pi \in GCWS(n-1) \Rightarrow \Pi$ is anonymous and efficient is easy and is already proved in Moulin [6]. So, it is sufficient to prove that every voting scheme in Moulin's class is weak reciprocate, and the characterization will be complete.

Let us take any voting scheme $\Pi \in GCWS(n-1)$; That is, we fix an arbitrary distribution of phantom voters $\alpha = (\alpha_1, \alpha_2, ..., \alpha_{n-1})$. We will prove that the median of the peaks and phantoms gives us a voting scheme that preserves the weak reciprocity property. The median is defined as follows: $m\left(p(R_1), p(R_2), ..., p(R_n), \alpha_1, ..., \alpha_{n-1}\right) \quad \Leftrightarrow \quad \#\left\{i \mid p(R_i) \leq m\right\} + \#\left\{i \mid \alpha_i \leq m\right\} \geq n$ and
$\#\left\{i \mid p(R_i) \geq m\right\} + \#\left\{i \mid \alpha_i \geq m\right\} \geq n$.

Suppose any fixed distribution of peaks $p(\mathbf{R}) = (p(R_1), p(R_2), ..., p(R_n))$, so that the social decision is $m = m(p(\mathbf{R}), \alpha)$ and that somebody -let us call him $i$- changes his peak -without loss of generality one whose peak is on the left of the median $(p(R_i) \leq m)$- to $p(R') \neq p(R_i)$ . The shift of the new choice will depend on the allocation of agent $i's$ new peak. There are two possibilities:

**1**- $p(R') \leq m(p(\mathbf{R}), \alpha)$ :

We need to know how this change will affect the remaining agents, in order to check the reciprocity of the voting scheme. Notice that in this case the cardinality of the set of agents at both sides of the initial median will not vary and the distribution of phantoms is always the same. By the definition of median, the new choice will be the same:

$$m(p(R', \mathbf{R}_{-i}), \alpha) = m' = m = m(p(\mathbf{R}), \alpha).$$

Every individual will be then indifferent with both distributions of peaks and, by weak reciprocity we should check that every agent with peak $p(R')$

should either not affect $i$ or improve $i's$ position. Let us consider any $j$ such that $p(R_j) \leq m$. If $j$ changes to $p(R')$, the total number of peaks on the left of the median will remain unchanged, so the median cannot vary: $m(p(R', \mathbf{R}_{-j}), \alpha) = m'' = m = m(p(\mathbf{R}), \alpha)$, so $i$ does not loose. Now, let us fix any individual $h$ with initial peak on the right of $m$ : If he changes his preferences to $x_i'$, the left side of the median increases its weight relative to the right side -which lose $j's$ vote- so, in the case of shifting the choice, it has to be to the left of the initial median, so it is true that: $m(p(R', \mathbf{R}_{-j}), \alpha) = m'' \leq m = m(p(\mathbf{R}), \alpha)$. But the change of just one individual cannot make the median jump over anybody's peak, so everybody with peaks strictly on the left of the initial median -including agent $i$- should gain with the change. The only remaining possibility is that of $p(R_i) = m$, but in this case, whenever $p(R') \leq m$, if the initial change changes the rule's choice, we are in case 2, and if it does not, nobody can individually make the decision shift to the left, so reciprocity holds in this case.

**2**- $p(R') \geq m$ :

In this case, it is easy to prove that $m' \geq m$ and every $p(R_j) > m$ implies $p(R_j) \geq m'$. By single-peakedness, $\forall j \in N$ such that $p(R_j) \geq m$, $m' R_j m$, and weak reciprocity: $m(p(R', \mathbf{R}_{-j}), \alpha) R_i m$ should hold. Notice that $m(p(R', \mathbf{R}_{-j}), \alpha) \leq m(p(R', \mathbf{R}_{-i}), \alpha)$, and by single-peakedness again it will always be true that

$m(p(R', \mathbf{R}_{-j}), \alpha) R_i m$. Let us see what happens with people on the left of the initial mean: For $p(R_j) \leq m$, everybody will be equal or worse off than before: $m R_j m(p(R', \mathbf{R}_{-i}), \alpha)$; so, by weak reciprocity we will expect $i$ to weakly lose if some $j$ such that $p(R_j) \leq m$ moves to $p(R')$. As $p(R') \geq m$, the following medians will coincide: $m(p(R', \mathbf{R}_{-j}), \alpha) = m(p(R', \mathbf{R}_{-i}), \alpha)$, and by single-peakedness -or simply looking at the definition of median above- $i$ will not improve his position and this holds: $m R_i m(p(R', \mathbf{R}_{-j}), \alpha)$. ■

The last result establishes the characterization of the large set of SCFs which are anonymous, efficient and weak reciprocate SCFs. As we said above, introspective solidarity interpreted as weak reciprocity allows for a larger set of procedures for making public decisions that the *effective* solidarity requirement represented by $WDUPR$. The important role given to the status quo when requiring the latter property along with efficiency disappears when we require weak reciprocity, so the SCF can be made much more sensitive and responsive to changes in the individuals' tastes.

Finally, it may be useful to comment the price we have to pay for this result with respect to that of strong reciprocity. We have yet argued that weak reciprocity is a weaker concept of solidarity than strong reciprocity, but it makes more sense, so that strong reciprocity is undoubtedly too stringent at a minimal conceptual cost. More interesting is the following question: since the efficiency and anonymity requirements are both needed to get the last theorem, we may wonder about what kind of weak reciprocate SCFs. are we eliminating by imposing anonymity and efficiency together. Since anonymity is implied by replacement monotonicity combined with efficiency, the efficiency property is the crucial assumption in order to compare both solidarity principles. We should then, expect both weak reciprocate and replacement monotonic SCFs. to exist outside the efficiency environment. The problem is that they may not be voting schemes and strategy-proof, so that the whole preference relations of the agents may be relevant to determine the outcome. This fact makes them too complex objects and difficult to implement. We can only provide the reader with two families of SCFs. of this kind that lay outside our analysis and they are anonymous, weak reciprocate and replacement monotonic, but they lack efficiency -in fact, they hardly select efficient alternatives and are manipulable-. The first class contains no voting scheme: Assume $\Re = \Re^{SP}$ and let us consider the family $\Psi = \{f^a \mid a \in [0, M]\}$ . Given $a \in [0, M]$ , let $f^a(\mathbf{R})$ be defined as: $\forall \mathbf{R} \in \Re,\ f^a(\mathbf{R}) =$

$$= \begin{cases} \arg\max\limits_{x} \bigcup\limits_{i \in N} \{x \in [0, M] \mid xI_i a\} & iff\ \ \forall i \in N,\ \#\{x \in [0, M] \mid xI_i a\} > 1 \\ M & \text{otherwise} \end{cases}$$

This function is not difficult to understand: it simply finds the largest point that is indifferent with the fixed one $a$ -or $M$ if there does not exist another one -for every individual- and then, selects the largest -the closest one to $M$-. Notice that this SCF makes broad use of the information outside the agents' peaks. Let us define

$$b_i(\mathbf{R}) = \begin{cases} \max\{\{x \mid xI_i a\}, a\} & iff\ \exists x \neq a\ s.t.\ xI_i a. \\ M & \text{otherwise} \end{cases} \quad \forall \mathbf{R} \in \Re.$$

Notice that any function in the class $\Psi$ is replacement monotonic since any shift in the function cannot jump over anybody's $b_i(\mathbf{R})$ so that either everybody gains or everybody looses. It is not strong reciprocate because whenever $f^a(\mathbf{R}) = b_j$, if agent $i$ with $b_i(\mathbf{R}) < b_j(\mathbf{R})$ changes to preferences such that $b_i'(\mathbf{R}) = a$, since $b_i'(\mathbf{R}) < b_i(\mathbf{R}) < b_j(\mathbf{R})$, the social choice does not change, and leave all the others indifferent, but whenever agent $j$ moves to $a$, the social choice shifts and everybody gains, so $i$ will not be indifferent.

Notice that weak reciprocity holds in any case.

The second class are voting schemes, and they are anonymous, weak reciprocate and replacement monotonic, but they are not efficient, strategy-proof and strong reciprocate. Consider the class $\Sigma = \{f^a \mid a \in [0, M]\}$. Given $a \in [0, M]$, let $f^a(\mathbf{R})$ be defined as: $\forall \mathbf{R} \in \Re$,

$$f^a(\mathbf{R}) = \begin{cases} p(R_1) & iff\ p(R_1) = p(R_2) = ... = p(R_n). \\ a & \text{otherwise} \end{cases}$$

## 2.4 Conclusions

We have investigated in this work the introspective solidarity principles of reciprocity in public goods environments when monetary compensations are not possible.

In a first step, we try to calibrate the power of the reciprocity property combined with anonymity in a general context with a finite set of alternatives, without imposing any domain restriction on the preference space. *Theorem 15* offers us a negative result. It is shown that we cannot find any anonymous and reciprocate SCF within this unrestricted domain. We are, then, compelled to impose some kind of structure on the space of preferences to obtain a positive result. In order to compare reciprocity with welfare-domination under preference-replacement, we move to the public good context with infinite alternatives defined into a closed interval on the real line, where the single-peakedness restriction is quite a natural assumption.

*Theorem 18* offers the answer within this new context and proves that there exist efficient, anonymous and weak reciprocate SCFs. Moreover, all of them are fully characterized and the class of functions that preserve both properties turns out to coincide with Moulin's class of *Generalized Condorcet winner solutions*. This result can be considered in two different ways. First, it is clear that we have achieved our goal of enlarging the small class of replacement monotonic SCFs by allowing for procedures that are more sensitive to individual preferences. Secondly and in a strategic context, we can consider the result as some kind of reinforcement of the class of strategy-proof SCFs. within the restricted domain of single-peakedness, since we show that they also satisfy some introspective solidarity principle.

*Theorem 17* explores the strong version of reciprocity in the public good context and concludes that when there are three or more individuals, there do not exist minimally responsive strong reciprocate and anonymous SCFs.

and they are not compatible with efficiency either. The reason why strong reciprocity is so much demanding than its weak version lies essentially on the treatment of changes in preferences that do not alter the social decision. Strong reciprocity is clearly overdemanding when it requires that when somebody changes and the social decision does not move, nobody else can make it shift with the other's preferences.

Whatever interpretation of the result we may like best, it may be worthwhile to point out the close relations between strategy-proofness and the reciprocity-anonymity condition in some restricted domains -not only that of single-peakedness, but that of strict orderings too-. When talking about reciprocate SCFs, we are imposing a fairness principle of equal treatment among individuals when someone suffers a preference mutation. The fairness principle in some of its usual forms may not make sense when people can *lie* about their real preferences. But notice that reciprocity is consistent even in this uncertain context of private information, and this conceptual consistency is obtained free from the implied strategy-proofness of the anonymous-reciprocate SCF.

# References


[1] Barberà, S. & B. Peleg (1990): "Strategy- Proof Voting Schemes with Continuous Preferences". Social Choice and Welfare 7, 31-38.

[2] Barberà, S., H. Sonnenschein & Zhou (1991): "Generalized Median Voter Schemes and Commitees". Journal of Economic Theory

[3] Barberà, S., & M. Jackson (1994): "A characterization of Strategy-Proof Social Choice Functions for Economies with Pure Public Goods". Social Choice and Welfare 11, 241-252.

[4] Ching, S., & W. Thomson (1993): "Population Monotonic Solutions in Public Good Economies with Single- peakedness. University of Rochester mimeo.

[5] Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result". Econometrica 41, 587-601.

[6] Moulin, H. (1980): "On Strategy- proofness and Single- peakedness". Public Choice 35, 437-455.

[7] Moulin, H. (1984): "Generalized Condorcet -winners for Single- peaked Preferences and Single- plateau Preferences". Social Choice and Welfare 1, 127-147.

[8] Moulin, H. (1987): "The Pure Compensation Problem: Egalitarian versus laissez- fairism". Quarterly Journal of Economics 102, 769-783.

[9] Satterthwaite, M.A. (1975): "Strategy- proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions". Journal of Economic Theory 10, 187-217.

[10] Schmeidler, D. & H. Sonnenschein (1978): "Two Proofs of the Gibbard-Satterthwaite Theorem on the Possibility of a Strategy- proof Social Choice Function". in Proceedings of a conference on decision theory and social ethics at Schloss Reisensburg, edited by H. Gottinger & W. Ensler. Dordrecht, Holland: Reidel.

[11] Thomson, W. (1983): "The Fair Division of a Fixed Supply among a Growing Population". Mathematics of Operations Research 8, 319-326.

[12] Thomson, W. (1990): "A Replacement Principle". University of Rochester mimeo.

[13] Thomson, W. (1993): "The Replacement Principle in Public Good Economies with Single- peaked Preferences". Economic Letters 42, 31-36.

[14] Thomson, W. (1994): "Resource- Monotonic Solutions to the Problem of Fair Division when Preferences are Single- peaked". Social Choice and Welfare 11, 205-223.

[15] Thomson, W. (1995a): "Population- Monotonic Solutions to the Problem of Fair Division when Preferences are Single- peaked". Economic Theory 5, 229-246.

[16] Thomson, W. (1995b): "Population- Monotonic Allocation Rules", in "Social Choice, Welfare and Ethics" (W.A. Barnett, H. Moulin, M. Salles and N. Schofield, Eds.), Chap. 4, pp. 79-124, Cambridge Univ. Press, Cambridge, UK, 1995.

[17] Thomson, W. (1997): "The Replacement Principle in Economies with Single- peaked Preferences". Journal of Economic Theory 76, 145-168.

[18] Thomson, W. (1998): "The Replacement Principle in Economies with Indivisible Goods". Social Choice and Welfare 15, 57-67.

[19] Thomson, W. (1999): "Welfare-Domination under Preference-Replacement; A Survey and Open Questions". Social Choice and Welfare, 16, 373-394.

# 3 A SOCIAL CHOICE TRADE OFF BETWEEN ALTERNATIVE FAIRNESS CONCEPTS: SOLIDARITY VERSUS RIGIDITY

## 3.1 Introduction

We have mentioned in the previous chapter that the *replacement principle* was introduced by Thomson [8] as a general requirement of solidarity among agents who jointly face a problem of fair allocation. In general terms, the principle requires that the consequences of changes in the parameters defining such a problem should affect agents in the same direction, except may be those whose characteristics also change along with the parameters: all others must gain, or all lose, relative to their situation before the change occurred.

The set of relevant parameters can vary from problem to problem. This gives rise to different axioms, each of which embodies the general requirement of solidarity for a specific environment. *Population monotonicity* requires that population growth without an increase of resources should not improve the share of those who were initially part of society. Here, the relevant parameters are those describing which agents belong to society at a given point in time. In contrast, *resource monotonicity* requires that when the set of agents does not change, but the overall amount of resources varies, then either all agents should lose, or all gain. In this paper we concentrate in a third version of the general principle of solidarity, which applies when the change in parameters involves a change in the preferences of some agent in society. If this entails a change in the allocation, then all agents whose preferences did not change should be affected in the same direction: either all of them should gain, or all lose. This axiom was first presented by Moulin [6] under the name of *replacement domination.* It has also been referred to as *replacement monotonicity* or *welfare domination under preference replacement* ($WDUPR$). Its consequences have been widely explored both for economies with private and public goods -Thomson [13], [14], [9] and [15].

When the problem faced by society is that of deciding the amount to provide of a pure public good, and the set of alternatives is viewed as an interval in the real line, Thomson [9] proved that, if the preferences of agents are single-peaked, the only efficient social choice functions satisfying $WDUPR$

are those in a narrow subclass within the family of *Generalized Condorcet winner solutions* defined by Moulin [4]. This subclass is characterized by a status quo value in the interval, which will prevail as the outcome as long as some agent's ideal is above the status quo value in the interval, and some other agent is below it. The outcome will only depart from the status quo whenever all agents unanimously agree than a lower level is desirable, or all would prefer a higher level. In these cases, the social outcome is the ideal of that agent whose peak is closest to the status quo. Hence, full satisfaction of the solidarity principle as expressed through the $WDUPR$ axiom is obtained at the cost of only admitting very rigid rules, which are barely responsive to the preferences of the different agents.

Rather than insisting in an absolute trade-off between solidarity and rigidity, we develop some measures of the degree of compliance of each one of these desirable features, and then use them to classify different social decision rules in terms of their respective abilities to satisfy each one of the two principles in different degrees -in the spirit of Campbell and Kelly's [2] Trade-off Theory. In particular, a partial order among the set of rules is proposed to be able to say what rule "unambiguously" dominates another in terms of solidarity and in terms of decisional "rigidity". The existence of a general trade-off is proved in that the rules that turn out to dominate any other in terms of solidarity -Thomson's class- are those that are dominated by any other in terms of rigidity and -to a large extent-, vice versa, the median voter rule being the only social choice function that is dominated by any other in terms of solidarity. When some stronger measures of both properties are defined to measure the "degree" of both solidarity and rigidity, a quantifiable trade-off emerges within the class of Generalized Condorcet winner solutions, that include both the median voter and Thomson's [9] rule. Basically, our last theorem proves that an appropriate generalization of the "qualified majority" of a voting rule can measure both the *solidarity degree* -the minimum number of voters that move in the same welfare direction when a voter changes her preferences- and the *rigidity degree*- the minimum number of voters than need to prefer a change of the social choice either to the right or to the left for this change to take place. Therefore, any rule within the Generalized Condorcet winner solutions can be classified by its solidarity-rigidity degree, so that given social preferences over the measures of both properties, a socially optimal "qualified majority" can be chosen in the constitutional stage.

The chapter proceeds as follows: in *Section 2*, the basic model and the definitions are established. *Sections 3* and 4 are devoted to the definitions and results regarding solidarity and rigidity respectively. *Section 5* deals with results in the restricted environment of Voting Schemes belonging to the family of Generalized Condorcet winner solutions and finally, we conclude with some comments.

## 3.2 The model

Consider a committee -the society- composed by a fixed finite set of *agents or individuals* $N = \{1, ..., n\}$, indexed by $i$, $j, h$ and $l$. The committee must choose the level of a public good -or the location of a public utility- in the interval $[0, M] \subset E$, where $E$ denotes de real line.

Every individual $i \in N$ is endowed with a complete preference pre-ordering $R_i$ -or *preference relation*- over the set of alternatives. The set of possible preferences is $\Re$. We denote by $P_i$ and $I_i$ the asymmetric and symmetric parts of $R_i$, standing for the strict and indifference relations associated with $R_i$.

The agents' preference relations are *continuous* and *single- peaked*. A preference relation $R_i$ on $[0, M]$ is single-peaked if there exits a unique number $p(R_i) \in [0, M]$ such that $\forall x, y \in [0, M]$, if $y < x \leq p(R_i)$ or $p(R_i) \leq x < y$, then, $xP_iy$. The number $p(R_i)$ is the *peak* of agent $i's$ preferences and it is obviously the most preferred alternative of agent $i$.

An ordered list of preference relations for all the individuals is a *preference profile* and will be denoted by $R = (R_i)_{i \in N} = (R_1, ..., R_n)$. We will frequently use the well- known notation: $R = (R_i, R_{-i})$ $\forall i \in N$. When preferences are single-peaked, the associated vector of peaks will be: $p(R) = (p(R_i))_{i \in N} \in [0, M]^n$.

Now, we model the social objectives. A *social choice function* (SCF) $f$ is a function which associates a chosen alternative to every preference profile and it will be denoted by $f : \Re^n \rightarrow [0, M]$.

We will be interested in a special class of SCFs, called *voting schemes*, which only use information about the agents' peaks. Hence, a voting scheme $\Pi$ is a social choice function for which the following holds:

$$\forall R, R' \in \Re^n \ s.t. \ p(R) = p(R') \Longrightarrow \Pi(R) = \Pi(R').$$

**Definition 19** *For each given $R \in \Re^n$, $x \in [0, M]$ is an* ***efficient alternative*** *if there is no $x' \in [0, M]$ with $x'R_ix$ $\forall i \in N$ and $x'P_ix$ for some $i \in N$. Let $P(R)$ be the set of efficient alternatives, given $R$.*

*A SCF $f$ is **efficient** if it selects efficient alternatives for each preference profile, i.e., $\forall R \in \Re^n,\ f(R) \in P(R)$.*

Since preferences $R_i \in \Re$ are single-peaked for all $i \in N$, it is easy to prove that $f$ is efficient whenever $\forall R \in \Re^n$,
$f(R) \in P(R) = [\min \{p(R_i) \ \mid i \in N\},\ \max \{p(R_i) \ \mid i \in N\}]$.

**Definition 20** *A SCF $f$ is **anonymous** if any permutation of the different values of its arguments yields the same alternative, i.e., if for all one-to-one mapping $\sigma : N \to N,\ f(R_1, ..., R_n) = f(R_{\sigma(1)}, ..., R_{\sigma(n)})\ \forall R \in \Re^n$.*

This property assures that no information about the individuals' names is used in the decision rule.

**Definition 21** *A SCF $f$ satisfies the property of **Welfare-domination under preference-replacement** (WDUPR)*[5] *if:*
$\forall i \in N,\ \forall R \in \Re^n,\ \forall R_i' \in \Re$, *then, either* $f(R)R_j f(R_i', R_{-i})\ \forall j \in N\backslash \{i\}$ *or* $f(R_i', R_{-i})R_j f(R)\ \forall j \in N\backslash \{i\}$ .

Any change in preferences of any individual move the welfare of the remaining agents in the same direction: either all of them gain or all of them lose -in the weak sense-.

**Definition 22** *A SCF $f$ is a **Generalized Condorcet Winner Solution-**$(n-1)$ $(GCWS(n-1))$ if $\exists \alpha = (\alpha_1, ..., \alpha_{n-1}) \in A^{n-1}$, called phantom voters or fixed ballots such that for all $R \in \Re^n$,*
$f(R) = m\,(p(R_1), p(R_2), ..., p(R_n), \alpha_1, ..., \alpha_{n-1})$, *where $m$ stands for the median*[6].

Moulin [4] proved that when preferences are single- peaked on the interval $[0, M]$ , the only anonymous, efficient and strategy- proof voting schemes on $[0, M]$ are those belonging to the family $GCWS(n-1)$.

[5] This property has also been called *replacement domination* and *replacement monotonicity.*

[6] The median is defined as:
$m\,(p(R_1), p(R_2), ..., p(R_n), \alpha_1, \alpha_2, ..., \alpha_{n-1}) \Leftrightarrow$
$\#\{i \mid p(R_i) \leq m\} + \#\{i \mid \alpha_i \leq m\} \geq n-1$ and
$\#\{i \mid p(R_i) \geq m\} + \#\{i \mid \alpha_i \geq m\} \geq n-1$. Moreover, we assume that the phantom voters are ordered such that $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_{n-1}$.

**Definition 23** *Suppose n is odd. We call the* ***Median Voter*** *SCF to* $f \in GCWS(n-1)$ *such that* $\alpha_1 = \alpha_2 = ... = \alpha_{\frac{n-1}{2}} = 0$ *and* $\alpha_{\frac{n-1}{2}+1} = ... = \alpha_{n-1} = M$. *(i.e., the median of the agents' revealed peaks).*

The following definition describes a family of solutions which only differ by one parameter $a \in [0, M]$ .This rule plays a central role in Thomson [9].

Basically, it will choose $a$ whenever it is efficient, and it will choose the peak of the agent who is closest to $a$, otherwise.

**Definition 24** *A SCF* $f^a$ *is* ***adjusted constant to*** $a \in [0, M]$ *if for all* $R \in \Re^n$,

$$f^a(R) = \begin{cases} a & if\ a \in P(R) \\ \min\{p(R_i) \mid i \in N\} & if\ a < \min\{p(R_i) \mid i \in N\} \\ \max\{p(R_i) \mid i \in N\} & if\ a > \max\{p(R_i) \mid i \in N\} \end{cases}$$

*Let us denote by* $\Phi$ *the family of adjusted constant SCFs; namely* $\Phi = \{f^a \mid a \in [0, M]\}$ *and* $f^a$ *is adjusted constant to* $a$.

Notice that all the SCFs within class $\Phi$ are anonymous and strategy-proof voting schemes and all of them belong to the family $GCWS(n-1)$ with the $n-1$ phantom voters located on the same point[7]

Thomson [9] proved that this subclass of the $GCWS(n-1)$ family is the set of efficient social choice functions satisfying $WDUPR$.

## 3.3 Solidarity

We now define two auxiliary functions which can be associated to any SCF $f$. They will be useful in discussing milder requirements than the solidarity axiom, but still in a similar spirit.

**Definition 25** *The* ***improvers*** *associated to SCF* $f$ *, preference profile* $R = (R_1, R_2, ..., R_n)$, *changed preferences* $R'$ *and agent* $i$ *is the function:* $I^f : \Re^{n+1} \times N \rightarrow \{0, 1, ..., n-1\}$ *defined as:*

$$I^f(R_1, R_2, ..., R_n, R', i) = \#\{j \in N \backslash \{i\}\ \text{such that}\ f(R_{-i}, R'_i) R_j f(R)\}.$$

The improvers set associated to some agents' change from a given distribution of peaks simply gives the number of agents who weakly gain with the change.

[7] That is, $\alpha_1 = \alpha_2 = ... = \alpha_{n-1} = a$.

**Definition 26** *The* ***losers*** *associated to SCF* $f$ *, preference profile* $R = (R_1, R_2, ..., R_n)$*, changed preferences* $R'$ *and agent* $i$ *is the function:* $L^f : \Re^{n+1} \times N \rightarrow \{0, 1, ..., n-1\}$ *defined as:*

$$L^f(R_1, R_2, ..., R_n, R', i) = \#\{j \in N \backslash \{i\} \text{ such that } f(R) R_j \ f(R_{-i}, R'_i)\}.$$

Conversely, the losers associated to some agents' change from a given preference profile is simply the number of agents who weakly lose with the change.

Note that, for every social choice function $f$: $\forall i \in N$, $\forall R \in \Re^n$, $\forall R' \in \Re$, it holds that $I^f(R_1, R_2, ..., R_n, R', i) + L^f(R_1, R_2, ..., R_n, R', i) + 1 \leq 2n - 1$.

We can now introduce a notion of the degree of solidarity associated to a preference profile and to the change in preferences of some agent embodied in a given social choice function.

**Definition 27** *The* ***solidarity degree*** *associated to SCF* $f$ *, agent* $i$ *, preference profile* $R = (R_1, R_2, ..., R_n)$*, when* $i$*'s preferences change to* $R'$ *is the function:* $S^f : \Re^{n+1} \times N \rightarrow \{0, 1, ..., n-1\}$ *defined as:*

$$S^f(R_1, R_2, ..., R_n, R', i) =$$
$$= \sup \left\{ I^f(R_1, R_2, ..., R_n, R', i),\ L^f(R_1, R_2, ..., R_n, R', i) \right\}.$$

The solidarity degree associated to each preference profile and any individual preference change is the maximum number of individuals who move together in the same welfare direction, as a result of that change. Following Thomson's general concept of solidarity, it seems natural to say that, given a change in somebody's preferences, a SCF with a greater solidarity degree than any other will behave better in solidarity terms. The problem that may arise when proposing a relaxation of the degree of solidarity is that a function might have a greater degree of solidarity than other for some profile and individual's change but a smaller degree of solidarity for some other profile and agent's preferences change, so in order to make solidarity judgements among different SCFs, society need to compare different solidarity behavior in different situations. One possibility is to classify every SCF in terms of solidarity using a pessimistic criterion: the minimum solidarity degree in every circumstance can be the *social utility level* -or the aggregate measure of solidarity- associated to a SCF, so that when society agree to use a SCF based in solidarity considerations follow a *maximin* rule on the solidarity degree functions of the admissible SCFs. Let us define the solidarity degree of a SCF:

**Definition 28** *We call **solidarity degree** associated to social choice function $f$ and denote as $SD^f$ to the following number:*

$$SD^f = \inf \left\{ S^f(R_1, ..., R_n, R', i) \ \mid (R_1, ..., R_n, R', i) \in \Re^{n+1} \times N \right\}.$$

The solidarity degree is the minimal number of individuals who move together in the same welfare direction when considering any possible configuration of peaks and any individual's change, so it only depends on the specific SCF, and each one will have associated just one solidarity degree. Note that every SCF in class $\Phi$ has the maximum possible solidarity degree, i.e., $SD^{f^a} = n - 1 \, \forall f^a \in \Phi$.

Nevertheless, we should recognize that there is no strong reason for such a pessimistic social preferences over SCFs. The only thing we can say about social preferences on solidarity is that if a SCF has a higher degree of solidarity than other -and strictly higher in some case- for every preference profile and individual change, the former will be unambiguously socially preferred to the latter when considering solidarity alone. Exploiting this *dominance* relation over SCFs in terms of solidarity, we can define the following:

**Definition 29** *We say that SCF $f$ is **dominated in terms of solidarity** by SCF $g$ if $\forall i \in N$, $\forall R \in \Re^n$, $\forall R' \in \Re$ such that $f(R) \neq f(R'_i, R_{-i})$,*

$$S^g(R_1, ..., R_n, R', i) \geq S^f(R_1, ..., R_n, R', i)$$

*and $\exists i \in N$, $\exists R \in \Re^n$, $\exists R' \in \Re$ such that $f(R) \neq f(R'_i, R_{-i})$ and*

$$S^g(R_1, ..., R_n, R', i) > S^f(R_1, ..., R_n, R', i).$$

*A SCF that is not dominated in terms of solidarity by those belonging to some class is said to be undominated within that class.*

Notice that the domination relation established above is incomplete and may generate cycles. Since we additionally impose the condition $f(R) \neq f(R'_i, R_{-i})$ in the above definition, it is possible that a SCF both dominates and is dominated by another. Restricting attention to efficient SCFs[8], there is a maximal set of SCFs from the above dominance relation, i.e., a set of functions that are not dominated by any other.

[8] We are implicitly assuming that efficiency is actually an essential admissibility requirement.

**Theorem 30** *The only efficient SCFs that are undominated in terms of solidarity among all efficient SCFs are the members of class* $\Phi$*.*

Actually, *Theorem 30* above is a straightforward reformulation of Thomson's [9] original result. Thomson proved that in this context, the only efficient SCFs such that $WDUPR$ holds are those belonging to class $\Phi$. However, notice that, as we said above, it amounts exactly to showing that the only efficient SCFs $f$ with $SD^f = n - 1$ are those in class $\Phi$. But any SCF exhibiting the highest solidarity degree has to be *the only* undominated SCFs among all the efficient SCFs.

The problem with the SCFs within class $\Phi$ is that they amount exactly to fix an alternative as a *status quo* that will only be changed with strict unanimity of all members of the committee, so the absolute solidarity is achieved at the price of high *decisional rigidity*, which can be considered as an additional fairness criterion and society will actually care about it.

We can now state the main results in the paper: *Theorems* 31, 37 and 40 motivate the paper and prove the existence of a trade-off between higher degrees of solidarity and low responsiveness -high rigidity- of SCFs. Moreover, the greatest degree of solidarity -Thomson's $WDUPR$- is associated to the least responsive SCFs -class $\Phi$- and the smallest possible solidarity level can only be satisfied by a much less rigid SCF in this context: the *Median Voter* SCF. The results lead us to focus in a broader class of voting schemes that include class $\Phi$ as well as the Median Voter SCF: the Generalized Condorcet winner solutions. The solidarity degree is then used to classify every SCF within this large class in *Theorem 40*, ranging from the one with the least solidarity degree -the median- to that with the highest solidarity degree -class $\Phi$-. Moreover, a direct trade-off is proved between the solidarity degree and the *rigidity degree* within that class.

Now, we prove our first characterization theorem, which can be viewed as a parallel to Thomson's [9] *Theorem 3.6* in the opposite side of the solidarity spectrum:

**Theorem 31** *Suppose that* $n > 3$ *is odd. The only efficient SCF that is dominated in terms of solidarity by all efficient SCFs is the Median Voter SCF.*

**Proof. Necessity**: Let us consider some SCF $f$, any preference profile $R \in \Re^n$ and suppose that some agent, say $i$, change his preferences to $R' \in \Re$

and the change shifts the social decision ($f(R) \neq f(R'_i, R_{-i})$); the smallest possible solidarity degree for the SCF facing such a change when $n$ is odd is obviously $S^f(R_1, ..., R_n, R', i) = \frac{n-1}{2}$, so if there would exist some set $C$ of efficient SCFs such that $\forall R \in \Re^n$, $\forall i \in N$, $\forall R' \in \Re$,

$$S^f(R_1, ..., R_n, R', i) = \frac{n-1}{2} \qquad \text{whenever} \quad f(R) \neq f(R'_i, R_{-i}) \qquad (1),$$

they will obviously be dominated in terms of solidarity by any other SCF not belonging to class $C$. We will show that the SCF $f$ such that $\forall R \in \Re^n$, $f(R) = m\{p(R_1), ..., p(R_n)\}$ is precisely the only efficient SCF in which (1) holds, so it is dominated by any other. Suppose any $R \in \Re^n$ and any efficient SCF $g$ such that (1) holds. Consider any preference profile $\overline{R} = (\overline{R}_1, \overline{R}_2, ..., \overline{R}_n)$ such that the following holds: $\forall i \in N$ such that $p(R_i) = m\{p(R_1), ..., p(R_n)\}$, $\overline{R}_i = R_i$ and $\forall i \in N$ such that $p(R_i) \neq m\{p(R_1), ..., p(R_n)\}$, $\overline{R}_i$ is any preference relation such that $p(\overline{R}_i) = m\{p(R_1), ..., p(R_n)\}$, so that the following holds: $p(\overline{R}_1) = ... = p(\overline{R}_i) = ... = p(\overline{R}_n) = m\{p(R_1), ..., p(R_n)\}$. By efficiency of $g$, it must be that $g(\overline{R}) = m\{p(R_1), ..., p(R_n)\}$. If $\overline{R} = R$, we are done. If $\overline{R} \neq R$, $\exists j \in N$ whose preferences in profile $\overline{R}$ have not the same peak of those in profile $R$. Let us take all these individuals and suppose w.l.g. that at least one of them is such that $p(R_j) < p(R_i)$ $\forall i \in N$ -if more than one, take any of them-, and consider the profile $(R_j, \overline{R}_{-j})$. Notice that it must be that $g(R_j, \overline{R}_{-j}) = g(\overline{R}) = m\{p(R_1), ..., p(R_n)\}$, since if not, the $n-1$ agents whose peaks lie on $g(\overline{R})$ in profile $(R_j, \overline{R}_{-j})$ would lose and therefore $\exists j \in N$, $\exists R_j \in \Re$, $\exists \overline{R} \in \Re$ with $g(\overline{R}) \neq g(R_j, \overline{R}_{-j})$ such that $S^g(\overline{R}_1, ..., \overline{R}_n, R', j) = n - 1 > \frac{n-1}{2}$. Now, consider profile $(R_j, \overline{R}_{-j})$. If $(R_j, \overline{R}_{-j}) = R$, we are done. If not, two possibilities can occur: **Case 1**. $\forall h \in N$, $p(\overline{R}_h) \leq m\{p(R_1), ..., p(R_n)\}$. In this case, we can iteratively change the preferences of all agents whose preferences in profile $\overline{R}$ do not have the same peaks as those in profile $R$ and the social choice cannot change, since $m\{p(R_1), ..., p(R_n)\}$ coincides with the greatest peak in profile $R$, so at least $\dfrac{n-1}{2} + 1$ individuals have their peaks in profile $R$ located in $m\{p(R_1), ..., p(R_n)\}$. Finally, we will construct profile $R$ from $\overline{R}$ and the social choice remains the median of the peaks in $R$, so we are done. **Case 2**. $\exists h \in N - \{j\}$ such that $p(R_h) > p(R_i)$ $\forall i \in N - \{j\}$ -if more than one, take any of them-. Take that individual and notice that it has to be true that $g(R_j, R_h, \overline{R}_{-j-h}) = g(R_j, \overline{R}_{-j}) = g(\overline{R}) = m\{p(R_1), ..., p(R_n)\}$, since if not, if $g(R_j, R_h, \overline{R}_{-j-h}) < g(R_j, \overline{R}_{-j})$, $S^g(R_j, \overline{R}_{-j}, R_h, h) \geq n-2 > \dfrac{n-1}{2}$ when $n > 3$ and if $g(R_j, \overline{R}_{-j}) < g(R_j, R_h, \overline{R}_{-j-h})$, $S^g(R_j, \overline{R}_{-j}, R_h, h) = n-1 > \dfrac{n-1}{2}$. Now, let us focus on profile $(R_j, R_h, \overline{R}_{-j-h})$ and consider any agent $l \in N - \{j, h\}$ such that $p(\overline{R}_l) \neq p(R_l)$. Again, two possibilities can occur: **Case1**. There does not exist such an agent. In this case, $(R_j, R_h, \overline{R}_{-j-h}) = R$ by construction of profile $\overline{R}$ from profile $R$, and therefore, we have shown that $m\{p(R_1), ..., p(R_n)\} = g(\overline{R}) = g(R_j, \overline{R}_{-j}) = g(R_j, R_h, \overline{R}_{-j-h}) = g(R)$ and we are done. **Case 2**. $\exists l \in N - \{j, h\}$ such that $p(\overline{R}_l) \neq p(R_l)$. In this case, consider agent's $l$ change from his preferences $\overline{R}_l$ in profile $(R_j, R_h, \overline{R}_{-j-h})$ to preferences $R_l$. if $g(R_j, R_h, R_l, \overline{R}_{-j-h-l}) \neq g(R_j, R_h, \overline{R}_{-j-h})$, then, $S^g(R_j, R_h, \overline{R}_{-j-h}, R_l, l) \geq \dfrac{n+1}{2} > \dfrac{n-1}{2}$, because, by construction of profile $\overline{R}$, it always holds for any individual change from preferences in $\overline{R}$ to preferences in $R$ that the median will be the same, i.e.,

$$m\left\{p(\overline{R})\right\} = m\left\{p(R_j, \overline{R}_{-j})\right\} = m\left\{p(R_j, R_h, \overline{R}_{-j-h})\right\} = \\ = m\left\{p(R_j, R_h, R_l, \overline{R}_{-j-h-l})\right\} = ... = m\{p(R)\}$$

, since, by definition of the median, it holds that $\forall S \subseteq \{i \in N \mid p(R_i) < m\{p(R)\}\}$ and $\forall T \subseteq \{i \in N \mid p(R_i) > m\{p(R)\}\}$

$$\#S + \#(N - \{S \cup T\}) \geq \#\{i \in N \mid p(R_i) \leq m\{p(R_1), ..., p(R_n)\}\} \geq \frac{n+1}{2}, \tag{4}$$

and

$$\#T + \#(N - \{S \cup T\}) \geq \#\{i \in N \mid p(R_i) \geq m\{p(R_1), ..., p(R_n)\}\} \geq \frac{n+1}{2}, \tag{5}$$

so if $g(R_j, R_h, R_l, \overline{R}_{-j-h-l}) \neq g(R_j, R_h, \overline{R}_{-j-h})$, $S^g(R_j, R_h, \overline{R}_{-j-h}, R_l, l) \geq \dfrac{n+1}{2} > \dfrac{n-1}{2}$ because the number of losers with $l$'s change (the number of agents whose peaks in profile $(R_j, R_h, \overline{R}_{-j-h})$ are on $m\{p(R)\}$ or on the left of $m\{p(R)\}$ if $g(R_j, R_h, R_l, \overline{R}_{-j-h-l}) > g(R_j, R_h, \overline{R}_{-j-h})$ or the number of agents whose peaks in the same profile are on the right of $m\{p(R)\}$ or exactly in that median if $g(R_j, R_h, R_l, \overline{R}_{-j-h-l}) < g(R_j, R_h, \overline{R}_{-j-h})$) cannot in any case be smaller than $\dfrac{n+1}{2}$. By construction of profile $\overline{R}$, we can reach profile $R$ by sequentially changing the preferences in profile $\overline{R}$ of one individual from the set $S$ or $T$ to his preferences in profile $R$ until exhausting them. When we move the last agent, and get a profile such that $\#S = \#T = \emptyset$ we get profile $R$. Since in no such change the social choice can ever shift from $m\{p(R)\}$, we get that necessarily, it must be that $g(R) = m\{p(R)\}$ and we can replicate the same argument for every $R \in \Re^n$, so the only efficient SCF with minimum solidarity degree is such that $\forall R \in \Re^n$, $g(R) = m\{p(R)\}$, i.e., the Median Voter SCF, which can also be written as the member of the class $GCWS(n-1)$ that allocates half of the phantom voters at each extreme of the interval -when $n$ is odd-.

**Sufficiency**: Let us consider the SCF such that $g(R) = m\{p(R)\}$ for any profile $R$ and notice that for any individual $i \in N$ such that $p(R_i) \leq m\{p(R)\}$, w.l.g, for all $R_i'$ such that $p(R_i') \leq m\{p(R)\}$, it holds that $g(R) = g(R_i', R_{-i})$, since the median cannot change. The only possible shift in the social choice comes from individual changes that *jump over the median*, i.e., $R_i'$ is such that $p(R_i') > m\{p(R)\}$ and just one individual different from $i$ in profile $R$ has the peak chosen by the median. In this case, the median shifts to the next peak on the right of the initial choice among the peaks in the vector $p(R_i', R_{-i})$, so that $\frac{n-1}{2}$ agents - those on the left of $m\{p(R)\}$ included the agent whose peak coincide with $m\{p(R)\}$ -with the exception of $i$- lose for sure and $\frac{n-1}{2}$ agents -those strictly on the right of $m\{p(R)\}$- gain for sure. A symmetric argument is used in the case that the individual who changes his preferences -say $i$- has an initial peak on the right of the median: $p(R_i) \geq m\{p(R)\}$. ■

Whenever $n > 3$ is odd, *Theorem 31* characterizes the only SCF which embodies the least solidarity in Thomson's sense, which also has the least solidarity degree. It is not actually needed to characterize the SCFs for every $n$ -the cases with just three agents or when the median is not defined-, since the logic of the proof makes the SCFs to approach the median. Notice that we come back to the Voting Schemes in the $GCWS(n-1)$ family, so the only efficient SCF with the smallest solidarity degree is additionally an *anonymous* and *strategy-proof* SCF, which is a generalization of simple majority voting for a unidimensional set of alternatives. Furthermore, *Theorem 31* suggests a new conjecture within voting schemes in the family of $GCWS(n-1)$: if the minimum solidarity degree of the **unanimity rules** -members of class $\Phi$- equals $n-1$ -*Theorem 30*- and the minimum solidarity degree of the simple majority voting rule -the median, provided that $n$ is odd- equals $\frac{n-1}{2}$ -

*Theorem 31*-, is it true that the minimum solidarity degree of a qualified majority voting rule amounts to be the required majority minus one?. It is easy to check that this conjecture is true. A qualified majority voting rule in this context can be described as resulting from an accumulation of phantom voters on the same point, when half of the remaining phantom voters are located in each of the extremes of the interval. The solidarity degree for any one of these SCFs equals the number of phantom voters in the accumulation point minus one -the required qualified majority-. This is not a surprising result, since a change in the rule due to any individual change in preferences can only take place when almost all the qualified majority of voters prefer other alternative to the status quo, for example, and a qualified majority implies that any change in the status quo should be supported by, at least, the required majority of voters, which will always gain with the change. Note, however, that these qualified majority voting rules do not exhaust the whole class of the $GCWS(n-1)$ family, since others which would be associated with a more disperse distribution of phantom voters could not be easily interpreted as a qualified majority to defeat a given status quo. Therefore, we know that the whole range of solidarity degrees are represented by some -efficient- SCF in the class of $GCWS(n-1)$, and the larger and smaller minimum solidarity degrees can only be found in that family.

## 3.4 Rigidity

Now we clarify what we understand by rigidity or relative flexibility of a SCF. We follow a similar approach to that used when defining solidarity. We need some more definitions: given SCF $f$ and any $R \in \Re^n$, let $FC^f(R)$ be the set of "coalitions" $S \subseteq N$ of agents that can impose a shift to the left on the social choice by changing all from their original preferences in profile $R$ to some other preferences $R'_S$ and all of them strictly gain with the change. In other words,

$$FC^f(R) = \left\{ \begin{array}{l} S \subseteq N \text{ such that } \exists R'_S \in \Re^S \text{ such that } f(R'_S, R_{-S}) < f(R) \\ \text{and } f(R')P'_i f(R) \; \forall i \in S \end{array} \right\}$$

To understand the intuition behind this property, consider an efficient SCF $f$ such that $0 < p(R_i) < M$ for all $i \in N$. It is clear that $N \in FC^f(R)$, since $\exists R'_i$ such that $p(R'_i) = 0$ for all $i \in N$ such that $0 = f(R') < f(R)$ and $f(R')P'_i f(R)$ for all $i \in N$. But to get a profitable change to the left maybe not every agent need to change her preferences, so other smaller coalitions may shift the SCF $f$ to the left.

**Definition 32** *The* ***Left coalition*** *associated to SCF $f$ and preference profile $R = (R_1, R_2, ..., R_n)$, is the function: $F^f : \Re^n \rightarrow \{1, 2, ..., n\}$ defined as: $\forall R \in \Re^n$, $F^f(R) = \inf_{S \subseteq N} \# \left\{ S \subseteq FC^f(R) \right\}$.*

Hence, this functions gives for each profile the *minimum* number of individuals that *have to prefer* a shift of the social decision to the left to be implemented. Notice that whenever $f(R) = 0$, $F^f(R) = \# \{\emptyset\} = 0$.

A symmetric argument gives us the coalition of agents that have to change preferences preferring a shift to the right for this to occur: given SCF $f$ and any $R \in \Re^n$,

$$GC^f(R) = \left\{ \begin{array}{l} S \subseteq N \text{ such that } \exists R'_S \in \Re^S \text{ such that } f(R'_S, R_{-S}) > f(R) \\ \text{and } f(R')P'_i f(R) \ \forall i \in S \end{array} \right\}$$

**Definition 33** *The* ***Right coalition*** *associated to SCF $f$ and preference profile $R = (R_1, R_2, ..., R_n)$, is the function: $G^f : \Re^n \rightarrow \{1, 2, ..., n\}$, defined as: $\forall R \in \Re^n$, $G^f(R) = \inf_{S \subseteq N} \# \left\{ S \subseteq GC^f(R) \right\}$.*

Again, this functions gives for each profile the *minimum* number of individuals that *have to prefer* a shift of the social decision to the right to take place. Notice that for all $R \in \Re^n$ such that $f(R) = M$, $G^f(R) = \# \{\emptyset\} = 0$.

**Definition 34** *We call* ***rigidity degree*** *associated to SCF $f$ and profile $R \in \Re^n$ and denote as $R^f$ to the following number:*

$$R^f(R) = \sup \left\{ F^f(R), \ G^f(R) \right\}.$$

The rigidity degree associated to a given profile is given by the greatest between the left and the right coalition for that profile. Since a SCF can show an asymmetric rigidity depending on the direction of the move -more propensity to shift to the left, or in other words, less supporters required for a change to the left, for example- we opt by taking the largest of them as a measure of rigidity given a profile. However, other rules of determining the degree of rigidity may be plausible, depending on the feelings of society and the choice of one or another is somehow arbitrary in essence. The rigidity degree is also a way to generalize the concept of qualified majority in two-issues voting procedures to the more general public goods economies.

**Definition 35** *We call* ***rigidity degree*** *associated to SCF $f$ and denote as $RD^f$ to the following number:*

$$RD^f = \inf \ \left\{ R^f(R_1, ..., R_n) \ \mid R \in \Re^n \right\}.$$

Again, like in the case of the solidarity degree of a SCF, a pessimistic criterion is proposed to get a single measure in terms of utility.

Now, we are in the conditions of defining a dominance relation based on our measure of rigidity for a profile.

**Definition 36** *We say that SCF $g$ is **dominated in terms of rigidity** by SCF $f$ if $\forall R \in \Re^n$,*

$$R^f(R_1, ..., R_n) \geq R^g(R_1, ..., R_n)$$

*and $\exists R \in \Re^n$ such that*

$$R^f(R_1, ..., R_n) > R^g(R_1, ..., R_n)$$

*A class of SCFs is undominated in terms of rigidity within a broader class if they are not dominated by any SCF within that broader class.*

The above dominance relation unambiguously compare different SCFs in terms of rigidity provided that society accepts the rigidity degree as a plausible measure of the rigidity of a SCF for any profile. Like in the case of solidarity, the above dominance relation is incomplete, but unlike the former, this one is transitive and cannot generate cycles. Now, a maximal class can be characterized among the efficient SCFs in terms of rigidity.

**Theorem 37** *The only efficient SCFs that are undominated in terms of rigidity by any other efficient SCF are the members of class $\Phi$.*

**Proof.** It is easy to check that $\forall f \in \Phi$, $RD^f = n$, so $f \in \Phi$ cannot be dominated by any other SCF. To show that they are the only efficient SCFs such that $RD^f = n$, let us assume that $f$ is an efficient SCF such that $RD^f = n$ (1). Choose any $x \in [0, M]$ and take any arbitrary $R \in \Re^n$ such that $p(R_i) = p(R_j) = x \ \forall i, j \in N$. Let us call $R(x)$ any such profile. Since $f$ is efficient, $f(R) = x$. First, we must note that for any two profiles $\overline{R}$ and $R'$ such that $p(\overline{R}_i) = p(\overline{R}_j) = p(R'_i) = p(R'_j) = x \ \forall i, j \in N$, $G^f(\overline{R}) = n \Rightarrow G^f(R') = n$ and $F^f(\overline{R}) = n \Rightarrow F^f(R') = n$.

Now, let us suppose, w.l.g. that $R^f(R) = G^f(R)$; let us move sequentially the preferences of the agents to any others $\overline{R}_1, \overline{R}_2, ..., \overline{R}_n$ such that $p(\overline{R}_i) = p(\overline{R}_j) > x \ \forall i, j \in N$. Then, by (1), only when the sequence finish, for profile $\overline{R}$, $f(\overline{R})$ can be different from $x$, and by efficiency again we know that $f(\overline{R}) = p(\overline{R}_i) = p(\overline{R}_j) > x \; \forall i,j \in N$. Moreover, by construction, $R^f(\overline{R}) = G^f(\overline{R})$, so fix an arbitrary agent $j \in N$ and move sequentially any others' preferences to any others such that $p(\widehat{R}_i) \geq f(\overline{R}) \; \forall i \neq j$. By (1), for any such profile, $f(\overline{R}_j, \widehat{R}_{-j}) = f(\overline{R}) = \min\left\{p(\overline{R}_j),\, p(\widehat{R}_i) \; \forall i \neq j\right\}$. Hence, we have proved that $\forall x \in [0, M]$ such that $R^f(R(x)) = G^f(R(x))$, $x' > x \Rightarrow R^f(R(x')) = G^f(R(x'))$ and $\forall \widetilde{R} \in \Re^n$ such that $p_i(\widetilde{R}_i) > x \; \forall i \in N$, $f(\widetilde{R}) = \min\left\{p(\widehat{R}_i) \; \forall i \in N\right\}$ (2). A symmetric argument gives us that $\forall x \in [0, M]$ such that $R^f(R(x)) = F^f(R(x))$, $x' < x \Rightarrow R^f(R(x')) = F^f(R(x'))$ and $\forall \widetilde{R} \in \Re^n$ such that $p(\widetilde{R}_i) < x$, $f(\widetilde{R}) = \max\left\{p(\widehat{R}_i) \; \forall i \in N\right\}$ (3). Now we define the sets:

$x^- = \left\{x \in [0, M] \mid R^f(R(x)) = F^f(R(x))\right\}$ and

$x^+ = \left\{x \in [0, M] \mid R^f(R(x)) = G^f(R(x))\right\}$. It is clear by (2) and (3) that $x^- \cap x^+$ has to be a singleton. Suppose not: $\exists x, x' \in x^- \cap x^+$, with $x < x'$ w.l.g. Then, $\exists R \in \Re^n$ with $x < p(R_1) < p(R_2) < ... < p(R_n) < x'$ such that both $f(R) = p(R_1)$ by (2) and $f(R) = p(R_n)$ by (3), a contradiction. Hence, $\exists \widehat{x} \in [0, M]$ (unique) such that $\forall R \in R^n$ such that $p(R_i) > \widehat{x} \; \forall i \in N \Rightarrow f(R) = \max\{p(R_i) \; \forall i \in N\}$ and $p(R_i) < \widehat{x} \; \forall i \in N \Rightarrow f(R) = \min\{p(R_i) \; \forall i \in N\}$. To conclude the proof, we have to show that for any other profile, the social choice is actually $\widehat{x}$, i.e., $\forall R' \in R^n$ such that $\exists i, j \in N$ such that $p(R'_i) > \widehat{x}$ and $p(R'_j) < \widehat{x} \Rightarrow f(R') = \widehat{x}$. Take any profile $R(\widehat{x})$ and change agent $i's$ preferences from $R_i(\widehat{x})$ to $R'_i$. Since $\widehat{x} = x^- \cap x^+$, (1) implies that $f(R'_i, R_{-i}(\widehat{x})) = f(R(\widehat{x})) = \widehat{x}$. Now change agent $j's$ preferences in profile $(R'_i, R_{-i}(\widehat{x}))$ to $R'_j$. Again, since $\widehat{x} = x^- \cap x^+$, (1) implies $f(R'_i, R'_j R_{-i-j}(\widehat{x})) = f(R'_i, R_{-i}(\widehat{x})) = f(R(\widehat{x})) = \widehat{x}$. Now, move sequentially any other agent preferences (for $h \neq i, j$) to $R'_h$, and by (1) again, it has to be that $f(R') = f(R(\widehat{x})) = \widehat{x}$, so necessarily $f \in \Phi$. ■

**Proposition 38** *If $n > 3$ is odd, the Median Voter SCF is such that $F^f(R) = G^f(R) = \dfrac{n+1}{2} \; \forall R \in \Re^n$, and hence has a rigidity degree equal to $\dfrac{n+1}{2}$, i.e., $RD^{MV} = \dfrac{n+1}{2}$.*

**Proof.** Easy to check. Moreover, the proposition is an straightforward corollary of *Theorem 37.* ■

**Corollary 39** *If $n > 3$ is odd, the Median Voter SCF is dominated by any other SCF $f \in \Phi$ in terms of rigidity.*

The above corollary is straightforward from *Proposition 38* and the definition of dominance in terms of rigidity.

Hence, the only efficient SCFs that are undominated in terms of both solidarity and rigidity are the members of class $\Phi$. Moreover, the only efficient SCF that performs the worst in terms of solidarity -the Median Voter SCF- possess an acceptable rigidity degree -much lower than those in family $\Phi$. Since society will in general prefer SCFs with as much solidarity as possible and as less rigidity as possible, a fundamental trade-off is proved to exist.

## 3.5 Voting Schemes

Once we have proved the existence of the trade-off, we will investigate it in detail, but unfortunately working with the whole set of efficient SCFs turns out to be technically very complicated when we depart from the maximal and minimal sets of solidarity and rigidity. From now on, we shall restrict attention to Voting Schemes belonging to the class of $GCWS(n-1)$, the reason being that both class $\Phi$ and the Median Voter SCF are included within that class and that this family is proved to be particularly interesting. Only those SCFs contained in $GCWS(n-1)$ are efficient, anonymous and strategy-proof -see Moulin [4], Barberà & Jackson [1]-, so the family is favored by both strategic properties (strategy-proofness) and other ethical requirements (anonymity). The question now is: can we classify all the members of the class of $GCWS(n-1)$ by their inherent solidarity-rigidity degrees?. In other words, can we find the voting schemes $\Pi$ belonging to the family of $GCWS(n-1)$ such that $SD^{\Pi}=k$ and $RD^{\Pi}=k$ holds?. As this subfamily of voting schemes is only parameterized by the vector of allocations of $n-1$ phantom voters -$\alpha$-, the question we are going to answer is simply which condition we have to impose on the distribution of phantom voters in order to $SD^{m(p(R),\alpha)}=k$ and $RD^{m(p(R),\alpha)}=k$. Since we are dealing with Voting Schemes, we will simplify notation considering that agents reveal only their *peaks*: $x=p(R)\in[0,M]^{n}$ . A vector of peaks will be $x=(x_S,x_{-S})\ \forall S\subseteq N$.

Given a Voting Scheme belonging to the class of $GCWS(n-1)$ and given any phantom voter $\alpha_i\in[0,M]$ , the *cumulative number of left phantoms* is the total number of phantom voters located in the same position or strictly on the left of $\alpha_i$.:

$$N(\alpha,\alpha_i)=\#\{j\in\{1,2,...,n-1\}\text{ such that }\alpha_j\leq\alpha_i\}$$

Notice that $\forall\alpha\in[0,M]^{n-1}$ , it will always be true that $N(\alpha,\alpha_{i+1})\geq$

$N(\alpha, \alpha_i)$ $\forall i \in \{1, ..., n-1\}$ and $N(\alpha, \alpha_{n-1}) = n-1$ . Similarly, we can define the symmetric concept:

Given a Voting Scheme belonging to the class of $GCWS(n-1)$ and given any phantom voter $\alpha_i \in [0, M]$, the *cumulative number of right phantoms* is the total number of phantom voters located in the same position or strictly on the right of $\alpha_i$.:

$$R(\alpha, \alpha_i) = \# \{j \in \{1, 2, ..., n-1\} \text{ such that } \alpha_j \geq \alpha_i\}.$$

Our main finding in this section is the following characterization result:

**Theorem 40** *Given a Voting Scheme* $m(x, \alpha) \in GCWS(n-1)$, $SD^{m(x,\alpha)} = RD^{m(x,\alpha)} - 1 = \inf_{\alpha_i} \sup \{N(\alpha, \alpha_i),\ (n-1) - N(\alpha, \alpha_i)\}$ .

**Proof.** We first prove[9] that $SD^{m(x,\alpha)} = RD^{m(x,\alpha)} - 1$. To see this, first notice that, since $m(x, \alpha) \in GCWS(n-1)$, for any $x \in [0, M]^n$, when only one agent changes his preferences -say $x_i' \neq x_i$-, the social choice cannot jump over anybody's peaks, so if $m(x_i', x_{-i}, \alpha) < m(x, \alpha)$, everybody with peaks to the left of $m(x_i', x_{-i}, \alpha)$ will be strictly better off and everybody with peaks to the right of $m(x, \alpha)$ will be strictly worse off, with nobody in the middle. Moreover, the only way for any agent to change the social choice to, say to the left, is by changing to preferences with peaks on the left of the initial social choice, and every individual who can individually change the social choice necessarily prefers the new choice with his new preferences. Now, $\forall x \in [0, M]^n$, let us define the two numbers:

$$\inf_{Q \subseteq N} \# \left\{ \begin{array}{c} Q \subseteq N \mid x_i \geq m(x, \alpha)\ \forall i \in Q,\ x_i' < m(x, \alpha) \text{ s.t.} \\ m(x_Q', x_{-Q}, \alpha) < m(x, \alpha) \end{array} \right\} = Q^-(x),$$

and

$$\inf_{Q \subseteq N} \# \left\{ \begin{array}{c} Q \subseteq N \mid x_i \leq m(x, \alpha)\ \forall i \in Q,\ x_i' > m(x, \alpha) \text{ s.t.} \\ m(x_Q', x_{-Q}, \alpha) > m(x, \alpha) \end{array} \right\} = Q^+(x).$$

The following holds for every $x \in [0, M]^n$ :

$F^{m(x,\alpha)}(x) =$

$$= \inf_{S \subseteq N} \# \left\{ \begin{array}{c} S \subseteq N \mid \exists x' \in [0, M]^n \text{ such that} \\ m(x', \alpha) < m(x, \alpha)\ \&\ m(x', \alpha) P_i' m(x, \alpha)\ \forall i \in S \end{array} \right\} = k \Leftrightarrow$$

$\Leftrightarrow Q^-(x) + \# \{i \in N \mid x_i \leq m(x, \alpha)\} = k$. Furthermore, for $G^{m(x,\alpha)}(x)$, we obtain:

[9] Since this proof takes any Voting Scheme $m(x, \alpha) \in GCWS(n-1)$ as given, we denote the function as $m(x, \alpha)$, $m(x)$ or just $m$.

$G^{m(x,\alpha)}(x) =$
$= \inf_{S\subseteq N} \# \left\{ \begin{array}{c} S \subseteq N \ \mid \exists x' \in [0,M]^n \text{ such that} \\ m(x',\alpha) > m(x,\alpha) \ \& \ m(x',\alpha) P_i' m(x,\alpha) \ \forall i \in S \end{array} \right\} = k \Leftrightarrow$
$\Leftrightarrow Q^+(x) + \# \{i \in N \mid x_i \geq m(x,\alpha)\} = k.$

We are proving that $RD^{m(x,\alpha)} = SD^{m(x,\alpha)} + 1$ and we proceed by contradiction:

**Step 1**: Let us suppose that $RD^{m(x,\alpha)} > SD^{m(x,\alpha)} + 1$. This means that $\exists x, x_i' \in [0,M]^{n+1}$ such that $SD^m(x,x_i') + 1 = \sup \{L^m(x,x_i'),\ I^m(x,x_i')\} + 1 < \sup \{F^m(x),\ G^m(x)\}\ \forall x \in [0,M]^n$ . Now, two cases can occur:

**Case 1**: $SD^m(x,x_i') = L^m(x,x_i')$. Consider the following vector of peaks: $\overline{x} = (x_i', x_{-i}) \in [0,M]^n$ . If $m(x_i', x_{-i}) < m(x)$, $G^m(\overline{x}) = L^m(x,x_i') + 1$ and it is easy to check that $F^m(\overline{x}) \geq I^m(x,x_i') + 1$ also because of the properties of the members of the family $GCWS(n-1)$, so we have found a vector $\overline{x} \in [0,M]^n$ such that $\sup \{F^m(\overline{x}),\ G^m(\overline{x})\} > SD^m(x,x_i') + 1$ and we are done. If $m(x_i', x_{-i}) > m(x)$, it happens that $F^m(\overline{x}) = L^m(x,x_i') + 1$ and $G^m(\overline{x}) \geq I^m(x,x_i') + 1$ and the result holds true in this case as well.

**Case 2**: $SD^m(x,x_i') = I^m(x,x_i')$. Consider the following vector of peaks: $\overline{x} = x \in [0,M]^n$ . If $m(x_i', x_{-i}) < m(x)$, $F^m(\overline{x}) = I^m(x,x_i') + 1$ and it is easy to check that $G^m(\overline{x}) \geq L^m(x,x_i') + 1$ so again we have found a vector $\overline{x} \in [0,M]^n$ such that $\sup \{F^m(\overline{x}),\ G^m(\overline{x})\} > SD^m(x,x_i') + 1$ and we are done. If $m(x_i', x_{-i}) > m(x)$, it occurs that $G^m(\overline{x}) = I^m(x,x_i') + 1$ and $F^m(\overline{x}) \geq L^m(x,x_i') + 1$ and in *Case 2* we are also able to find that the assumption in *Step 1* cannot be true, so given any $m(x,\alpha) \in GCWS(n-1)$, $RD^{m(x,\alpha)} \leq SD^{m(x,\alpha)} + 1$.

**Step 2**: Let us suppose that $RD^{m(x,\alpha)} < SD^{m(x,\alpha)} + 1$. This means that $\exists \overline{x} \in [0,M]^n$ such that $\forall x, x_i' \in [0,M]^{n+1}$ ,

$SD^m(x,x_i') + 1 = \sup \{L^m(x,x_i'),\ I^m(x,x_i')\} + 1 > \sup \{F^m(\overline{x}),\ G^m(\overline{x})\}$ . Now, two cases can occur:

**Case 1**: Let us suppose that $R^m(\overline{x}) = \sup \{F^m(\overline{x}),\ G^m(\overline{x})\} = F^m(\overline{x})$ and take any sequence of shifts of peaks such that $x_j \geq m(\overline{x},\alpha)$ to the left hand side of $m(\overline{x},\alpha) : x_j' < m(\overline{x},\alpha)$ until one more change shifts the social decision to $m(x'_{S\cup\{i\}}, \overline{x}_{-S}, \alpha) < m(\overline{x},\alpha)$ and $m(x'_S, \overline{x}_{-S}, \alpha) = m(\overline{x},\alpha)$. Then, $\exists x, x_i' = (x'_S, \overline{x}_{-S}, x_i') \in [0,M]^{n+1}$ such that $I^m(x'_S, \overline{x}_{-S}, x_i') = F^m(\overline{x}) - 1$ and $L^m(x'_S, \overline{x}_{-S}, x_i') \leq G^m(\overline{x}) - 1$, entering into a contradiction with the assumption in *Step 2*.

**Case 2**: Let us suppose now that $R^m(\overline{x}) = \sup \{F^m(\overline{x}),\ G^m(\overline{x})\} = G^m(\overline{x})$ and take any sequence of peaks such that $x_j \leq m(\overline{x},\alpha)$ to the right hand side of $m(\overline{x},\alpha) : x'_j > m(\overline{x},\alpha)$ until the last peak changed shifts the choice to $m(x'_{S\cup\{i\}},\overline{x}_{-S},\alpha) > m(\overline{x},\alpha)$ and $m(x'_S,\overline{x}_{-S},\alpha) = m(\overline{x},\alpha)$. Now, the same argument in *Case 1* applies now to the right: $\exists x, x'_i = (x'_S,\overline{x}_{-S},x'_i) \in [0,M]^{n+1}$ such that $I^m(x'_S,\overline{x}_{-S},x'_i) = G^m(\overline{x}) - 1$ and $L^m(x'_S,\overline{x}_{-S},x'_i) \leq F^m(\overline{x}) - 1$, so the assumption motivating *Step 2* cannot be true and together *Steps 1* and *2* imply the first part of the inequality in *Theorem 40*.

Now, we prove that $SD^{m(x,\alpha)}$ coincide with the last part of the equation in *Theorem 40*.

The function can also be written as:

$SD(m(x,\alpha)) = \inf\limits_{x,x'_i} \sup \{I(x,x'_i), L(x,x'_i)\} =$

$= \inf\limits_{\alpha_i} \sup \{N(\alpha,\alpha_i),\ (n-1) - N(\alpha,\alpha_i)\}$. We will need some lemmata:

**Lemma 41** *If* $m = m(x,\alpha),\ \forall x, x'_i \in [0,M]^{n+1}$ *such that* $\# \{i \mid x_i \leq m\} + N(\alpha,m) > n$ *and* $\# \{i \mid x_i \geq m\} + R(\alpha,m) > n \Rightarrow$

$\Rightarrow \sup \{I(x,x'_i),\ L(x,x'_i)\} = n-1.$

**Proof.** The median can be defined as follows: $m(x,\alpha) = m \Leftrightarrow$

$\Leftrightarrow \# \{i \mid x_i \leq m\} + N(\alpha,m) \geq n$ and $\# \{i \mid x_i \geq m\} + R(\alpha,m) \geq n$. Now, take any $x$ and for all $x'_i$ it will always hold: $\# \{i \mid x_i \leq m\} + N(\alpha,m) \geq n+1$ and $\# \{i \mid x_i \geq m\} + R(\alpha,m) \geq n+1 \Leftrightarrow m = m'$.

It is straightforward that if we subtract or add one unit to each side of the above expressions, we can represent every change in the allocation of any peak. Furthermore, by definition of $I(x,x'_i)$ and $L(x,x'_i)$ and since $m = m'$, we can write: $I(x,x'_i) = \# \{j \in N\backslash \{i\} \mid m(x_{-i},x'_i,\alpha) R_j m(x,\alpha)\} =$

$= \# \{j \in N\backslash \{i\} \mid m' R_j m\} = \# \{j \in N\backslash \{i\} \mid m R_j m\} = n-1$, and the same is true for $L(x,x'_i)$.

Hence, we have: $I(x,x'_i) = n-1$ and $L(x,x'_i) = n-1$, which implies: $S(x,x'_i) = \sup \{I(x,x'_i),\ L(x,x'_i)\} = \sup \{n-1,\ n-1\} = n-1$. ■

**Lemma 42** *First, let us define two given values to avoid large expressions:* $\alpha_0 = 0$ *and* $\alpha_n = M$. *Values which should not be confused with the phantoms* $\alpha' s$.

$\forall x \in [0,M]^n,\ \forall x'_i \in [0,M],$

$$\inf_{\substack{x,x'_i \\ x \in [\alpha_h,\alpha_t]}} \sup \{I(x,x'_i),\ L(x,x'_i)\} \geq \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \inf_{\substack{x,x'_i \\ x \in [\alpha_i,\alpha_{i+1}]}} \sup \{I(x,x'_i),\ L(x,x'_i)\}$$

**Proof.** Let us take an arbitrary $x \in [0, M]^n$, $x'_i \in [0, M]$ and construct another from it, defined as follows: $y = (y_1, y_2, ..., y_n)$, such that: $\forall i \in N,\ y_i = \inf(m, m') + [\sup(m, m') - \inf(m, m')] \dfrac{i}{n+1}$,

where $m = m(x, \alpha)$ and $m' = m(x, x'_i, \alpha)$.

Let us call: $\overline{\alpha} = \min_i \{\alpha_i \mid \alpha_i \geq m, m'\}$ and $\overline{\beta} = \max_i \{\alpha_i \mid \alpha_i \leq m, m'\}$.

Then, it is easy to see that: $\forall i \in N,\ y_i \in (\overline{\beta}, \overline{\alpha})$ ( open interval ). What we have done with the construction is to partition the interval $[m, m']$ into $n+1$ pieces and, whenever $m \neq m'$, it holds that $\forall i, j \in N,\ y_i \neq y_j$.

Now, let us call $\overline{m} = m(y, \alpha)$. Notice that the following will hold:

$N(\alpha, m) = N(\alpha, \overline{m})$ and $R(\alpha, m) = R(\alpha, \overline{m})$ (1). Furthermore, it will be true that: $\exists j \in N$ such that $y_j = \overline{m}$ and: $\#\{i \mid y_i \leq \overline{m}\} + N(\alpha, \overline{m}) = n$ and $\#\{i \mid y_i \geq \overline{m}\} + R(\alpha, \overline{m}) = n$ (2).

These last expressions mean that the properties that define the median hold with equality, and implies: $\sup\{I(y, y'_i, \alpha),\ L(y, y'_i, \alpha)\} \leq n - 1,\ \forall y \in [0, M]^n,\ \forall y'_i \in [0, M]$. Now, two things can happen:

**1-** $\#\{i \mid x_i \leq m\} + N(\alpha, m) > n$ and $\#\{i \mid x_i \geq m\} + R(\alpha, m) > n$, in which case, by *Lemma 41*: $\sup\{I(x, x'_i),\ L(x, x'_i)\} = n - 1$. And:

$\sup\{I(y, y'_i, \alpha),\ L(y, y'_i, \alpha)\} \leq n - 1 = \sup\{I(x, x'_i),\ L(x, x'_i)\}$ and this is true $\forall x \in [0, M]^n,\ \forall x'_i \in [0, M]$, and, in particular:

$$\inf_{x, x'_i, x \in [\alpha_h, \alpha_t]} \sup\{I(x, x'_i),\ L(x, x'_i)\} \geq \inf_{x, x'_i} \sup\{I(x, x'_i),\ L(x, x'_i)\}$$

**2-** $\#\{i \mid x_i \leq m\} + N(\alpha, m) = n$ and $\#\{i \mid x_i \geq m\} + R(\alpha, m) = n$, But we know that (1) and (2) hold, so: $\#\{i \mid x_i \leq m\} = \#\{i \mid y_i \leq m\}$ and $\#\{i \mid x_i \geq m\} = \#\{i \mid y_i \geq m\}$, and it is easy to prove that, by *Lemma 41*: $\forall x \in [0, M]^n,\ \forall x'_i \in [0, M]$ with $m \neq m'$ and $\exists j \in N$ such that $x_j = m \Rightarrow \{I(x, x'_i),\ L(x, x'_i)\} = \{\#\{i \mid x_i \leq m\} - 1,\ \#\{i \mid x_i \geq m\} - 1\}$, and, if $\exists \alpha_j$ such that $\alpha_j = m$ with $m \neq m'$, it is easy to check that:

$\sup\{I(x, x'_i),\ L(x, x'_i)\} \geq \{\#\{i \mid x_i \leq m\} - 1,\ \#\{i \mid x_i \geq m\} - 1\}$, and this implies:

$$\sup\{I(y, y'_i, \alpha),\ L(y, y'_i, \alpha)\} \leq \sup\{I(x, x'_i, \alpha),\ L(x, x'_i, \alpha)\}$$

$\forall x_i' \in [0, M]$, $\forall y_i' \in [0, M]$. We have proved, then, that the following statement is true:

$$\inf_{\substack{x,x_i' \\ x \in [\alpha_h, \alpha_t]}} \sup \{I(x, x_i'),\ L(x, x_i')\} \geq \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \inf_{\substack{x,x_i' \\ x \in [\alpha_i, \alpha_{i+1}]}} \sup \{I(x, x_i'),\ L(x, x_i')\}$$

$\forall h, t \in N$, because $\forall i \in N,\ y_i \in \left(\overline{\beta}, \overline{\alpha}\right)$ and there does not exist $\alpha_j \in \left(\overline{\beta}, \overline{\alpha}\right)$.
■

Coming back to the main proposition, and recalling that if we have a finite and countable set of real numbers $K$ and divide it into $m$ arbitrary subsets - indexed by $i$ -, called $K_i$, if $x_j \in E$ is a typical element of subset $K_j \subset K$, it always hold that $\inf \{x \in K\} = \inf_i \{\inf \{x_i \in K_i\}\}$. But the expression we are trying to prove was:

$$SD^{m(x,\alpha)} =$$

$$\inf_{\alpha_i} \sup \{N(\alpha, \alpha_i),\ (n-1) - N(\alpha, \alpha_i)\} = \inf_{\substack{h,t \\ h,t \in [0,1,...,n]}} \inf_{\substack{x,x_i' \\ x \in [\alpha_h, \alpha_t]}} \sup \{I(x, x_i')),\ L(x, x_i')\}$$

And this is true by *Lemma 42* and because $[\alpha_0, \alpha_n] = \bigcup_{h,t} [\alpha_h, \alpha_t]$. This last statement can be written as: $\inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \inf_{\substack{x,x_i' \\ x \in [\alpha_i, \alpha_{i+1}]}} \sup \{I(x, x_i'),\ L(x, x_i')\}$. Furthermore, as we said above, the last expression coincides with ( for $y's$ peaks profile of every interval without phantom voters):

$$\inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \inf_{\substack{x,x_i' \\ x \in [\alpha_i, \alpha_{i+1}]}} \sup \{\# \{i \mid x_i \leq m\} - 1,\ \# \{i \mid x_i \geq m\} - 1\} = (*)$$

By definition of the median in this interval: $\# \{i \mid x_i \leq m\} + N(\alpha, m) = n$ and $\# \{i \mid x_i \geq m\} + R(\alpha, m) = n$, and, since we are in an interval with no phantom voters in it: $N(\alpha, m) + R(\alpha, m) = n - 1 \Rightarrow R(\alpha, m) = n - 1 - N(\alpha, m)$, and substituting above $(*)$, we have:

$$\inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{n - N(\alpha, m) - 1,\ n - R(\alpha, m) - 1\} =$$

$$= \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{n - N(\alpha, m) - 1,\ n - (n - 1 - N(\alpha, m)) - 1\} =$$

$$= \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{N(\alpha, m),\ R(\alpha, m)\} = \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{N(\alpha, m),\ (n-1) - N(\alpha, m)\} =$$

$= \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{N(\alpha, \alpha_i),\ (n-1) - N(\alpha, \alpha_i)\}$. This last step is justified because: $N(\alpha, m) = N(\alpha, \overline{m}) = N(\alpha, \alpha_i)$ for some $i$ and $R(\alpha, m) = R(\alpha, \overline{m}) = R(\alpha, \alpha_i)$ for some $i$ (1). ■

The strategy of this proof is not difficult: we have shown that, given any vector of peaks $x$ in the whole interval $[0, M]$, if any individual change his peak to any other peak $x_i'$, the solidarity degree cannot become smaller with respect to a distribution of peaks inside some pair of contiguous phantoms. *Theorem 40* tells us that the solidarity degree we can expect from a *Generalized Condorcet winner solution* can be obtained this way: For any two different phantom voters location, choose the supreme between the cumulative number of phantoms at each side of the two extremes and then, take the minimum of all of them.

It is interesting to remark that Thomson's solution for the solidarity degree $n-1$ -$WDUPR$- is a particular case of functions such that the $SD^f = n-1$ and it is easy to see that the only voting schemes belonging to $GCWS(n-1)$ -in fact, the only SCFs- such that welfare-domination under preference-replacement -joint with efficiency- hold ($SD^f = n-1$) are those characterized by Thomson: The only way to get $SD^f = n-1$ with the above restriction on the phantom's distributions is to allocate all the $n-1$ phantom voters on a given point of the interval, which will ensure that: $\inf_{\alpha_i} \sup \{N(\alpha, \alpha_i),\ (n-1) - N(\alpha, \alpha_i)\} = n-1$. with $i \in \{0, 1, ..., n-1\}$.

Because there is only one allocation of phantoms, this expression can be written: $\inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{N(\alpha, \alpha_i),\ (n-1) - N(\alpha, \alpha_i)\}. = \inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{n-1,\ 0\} = n-1.$

Although Thomson's family $\Phi$ comes out when requiring $WDUPR$, the median is not the only SCF in the class of $GCWS(n-1)$ that has a solidarity degree of $\frac{n-1}{2}$ when $n$ is odd, since the maximin criterion assumed in the definition of the solidarity degree is less restrictive that requiring a SCF be dominated by any other (i.e., requiring that *for all* changes in the rule due to any individual's change, the number of losers and winners will be the same). Actually, there are a lot of SCFs such that $SD^f = \frac{n-1}{2}$ -i.e., all of those that allocate the $n-1$ phantom voters in different places- and, of course, dominate the Median Voter SCF in terms of solidarity.

*Theorem 40* offers a feasibility constraint faced by society in its constitutional stage -when deciding the SCF to use to select alternatives-. But when the ethical properties society cares about are measurable and society agrees with the measures used, society may have different sensibility about the degrees of fulfillment of the different ethical properties. This preferences defined over partial fulfillment of desirable properties of SCFs can be represented by a social utility function defined on them. Let $E$ be the real line and $C$ be a class of SCFs under consideration. Given a set of $k \geq 2$ measurable properties $X_i : C \rightarrow E, \ \forall i \in \{1, ..., k\}$, we can define the following:

**Definition 43** *A* ***Constitutional Social Welfare Function*** *is a function* $CW : \prod_{i=1}^{k} X_i \rightarrow E.$

**Definition 44** *A SCF $f \in C$ is* ***socially optimal at the constitutional stage*** *if* $CW(X_1(f), X_2(f), ..., X_k(f)) \geq CW(X_1(g), X_2(g), ..., X_k(g)) \ \forall g \in C.$

We can easily obtain results regarding social optimality at the constitutional stage once we know the nature of the trade-off between different properties. Applying the general setup to our problem and using *Theorem 40*, we can easily obtain different results that take the trade-off into account.

**Corollary 45** *Given* $C = GCWS(n-1), \ k = 1, 2, \ X_1 = SD(f), \ X_2 = RD(f)$ *and* $CW = SD - RD$, *every* $f \in GCWS(n-1)$ *is socially optimal at the constitutional stage.*

**Corollary 46** *Given* $C = GCWS(n-1), \ k = 1, 2, \ X_1 = SD(f), \ X_2 = RD(f)$ *and either* $CW = SD(n - RD)$ *or* $CW = \min\{SD, \ (n - RD)\}$, *the Median Voter rule is the only SCF that is socially optimal at the constitutional stage.*

**Corollary 47** *Given* $C = GCWS(n-1), \ k = 1, 2, \ X_1 = SD(f), \ X_2 = RD(f)$ *and either* $CW = \left(SD - \frac{n-1}{2}\right)(n - RD)$ *or*

$CW = \min\left\{\left(SD - \frac{n-1}{2}\right), \ (n - RD)\right\}$, *if* $\frac{3}{4}(n-1)$ *is an integer, the voting schemes such that*

$$\inf_{\substack{\alpha_i \\ i \in \{0,1,...,n-1\}}} \sup \{N(\alpha, \alpha_i), \ (n-1) - N(\alpha, \alpha_i)\} = \frac{3}{4}(n-1)$$ *are the only* $f \in C$ *that are socially optimal at the constitutional stage.*

The above approach has been used in the inequality measures literature and in the Trade-off theory approach to Social Choice Theory due to Campbell and Kelly [2]. The above corollaries are straightforward given *Theorem 40.*

## 3.6 Conclusions

We have investigated in this chapter a direct relaxation of Thomson's welfare-domination under preference-replacement property that allows for the utility of a given number of agents to move differently than that of others when some agent changes his preferences. The solidarity degree associated to a social choice function (SCF) is an index that allow us to classify different SCFs according to the degree of solidarity they exhibit. Thomson's family $\Phi$ will appear when requiring the maximum possible solidarity degree. We prove that, provided that the number of individuals is odd and not less than 4, there is a unique SCF that is efficient and is dominated in terms of solidarity by any other -it has the smallest solidarity degree possible in every circumstance according to our classification-. This SCF is the median of the agents' revealed peaks, and it can be considered a much less rigid SCF than those in family $\Phi$. Moreover, members of class $\Phi$ are proved to be the only efficient and undominated SCFs in terms of rigidity among all efficient SCFs. Hence, a basic trade-off between solidarity and rigidity is pointed out: more solidarity can only be obtained at the expense of less flexibility of the SCFs. SCFs covering the whole range of solidarity degrees can be found within a specially interesting class of voting procedures, that of Generalized Condorcet winner solutions. Therefore, we concentrate on studying the solidarity behavior of those Voting Schemes within this restricted context of the $GCWS(n-1)$ family. A complete description of such functions is provided in the last section. *Theorem 40* gives us a simple condition that any voting scheme belonging to $GCWS(n-1)$ should respect for having a specific minimum solidarity and rigidity degree. This condition can be viewed as a filter that allow us to classify every SCF of the family according to different solidarity degrees -in Thomson's sense- and different rigidity degrees, obtaining a *solidarity-rigidity concepts menu* which may help in the understanding of the fairness of social decision procedures at their constitutional stage. This paper can be also understood as an application of Campbell and Kelly's [2] view of Social Choice when different properties of SCFs can be relaxed and measured.

# References


[1] Barberà, S., & M. Jackson (1994): "A characterization of Strategy-Proof Social Choice Functions for Economies with Pure Public Goods". Social Choice and Welfare 11, 241-252.

[2] Campbell, D.E. & Kelly, J.S.:"$t$ or $1 - t$: That is the Tradeoff", Econometrica 61, (1993), 1355-1365.

[3] Ching, S., & W. Thomson (1993): "Population Monotonic Solutions in Public Good Economies with Single- peakedness. University of Rochester mimeo.

[4] Moulin, H. (1980): "On Strategy- proofness and Single- peakedness". Public Choice 35, 437-455.

[5] Moulin, H. (1984): "Generalized Condorcet -winners for Single- peaked Preferences and Single- plateau Preferences". Social Choice and Welfare 1, 127-147.

[6] Moulin, H. (1987): "The Pure Compensation Problem: Egalitarian versus laissez-fairism". Quarterly Journal of Economics 102, 769-783.

[7] Thomson, W. (1983): "The Fair Division of a Fixed Supply among a Growing Population". Mathematics of Operations Research 8, 319-326.

[8] Thomson, W. (1990): "A Replacement Principle". University of Rochester mimeo.

[9] Thomson, W. (1993): "The Replacement Principle in Public Good Economies with Single- peaked Preferences". Economic Letters 42, 31-36.

[10] Thomson, W. (1994): "Resource- Monotonic Solutions to the Problem of Fair Division when Preferences are Single- peaked". Social Choice and Welfare 11, 205-223.

[11] Thomson, W. (1995a): "Population- Monotonic Solutions to the Problem of Fair Division when Preferences are Single- peaked". Economic Theory 5, 229-246.

[12] Thomson, W. (1995b): "Population- Monotonic Allocation Rules", in "Social Choice, Welfare and Ethics" (W.A. Barnett, H. Moulin, M. Salles and N. Schofield, Eds.), Chap. 4, pp. 79-124, Cambridge Univ. Press, Cambridge, UK, 1995.

[13] Thomson, W. (1997): "The Replacement Principle in Economies with Single- peaked Preferences". Journal of Economic Theory 76, 145-168.

[14] Thomson, W. (1998): "The Replacement Principle in Economies with Indivisible Goods". Social Choice and Welfare 15, 57-67.

[15] Thomson, W. (1999): "Welfare-Domination under Preference-Replacement; A Survey and Open Questions". Social Choice and Welfare 16, 373-394.

# 4 DOMINANT STRATEGIES IMPLEMENTATION WHEN COMPENSATIONS ARE ALLOWED: A CHARACTERIZATION

## 4.1 Introduction

Since the early 70's, the problem of designing suitable incentive mechanisms to achieve socially desirable alternatives has been a major concern in economics. The initial negative results due to Gibbard [4] and Satterthwaite [14] in the context of unrestricted domains of preferences proved the need to impose domain restrictions to find some possibility results -see Dasgupta et al. [3] for a survey-. The first successful attempt to find a possibility result in mixed economies -those combining some public good with a private one- were due to Groves [6], [7] and [8], Clarke [2] and Green & Laffont [5].

We are interested in testing a regularity that emerges in many results regarding implementability in dominant strategies of social choice rules. In order to motivate our approach, let us consider three well-known results:

1. *The Gibbard-Satterthwaite Theorem* (Gibbard [4], Satterthwaite [14])

Let us consider a society in which a finite number $n \geq 2$ of agents or individuals, ordered in a set $N = \{1, ..., n\}$ and indexed by $i, j \in N$ choose **alternatives**, social states or objects from some set $K$, $\#K > 1$. These objects may be levels of provision of some public or private goods, allocations of indivisible goods, candidates for being a boss, etc.-. Let $k \in K$ denote any alternative from that set. Each individual is endowed with some private characteristic or **type** $\theta_i$ from a set $\Theta_i$. A **profile** is an element in the Cartesian product of the sets $\Theta_i$ $\forall i \in N$. Society can be described by a possible profile $\theta = (\theta_1, ..., \theta_n) = (\theta_i, \theta_{-i}) \in \prod_{i=1}^{n} \Theta_i$. We call here an **economy** to the tuple $e = \langle N, K, \Theta_i \forall i \in N \rangle$ . Given an economy, society -or the social planner- would like to select certain alternatives depending on which are the individual characteristics, so social desirability is summarized by a choice rule denoted by $K^* : \prod_{i=1}^{n} \Theta_i \to\to K$, called **social choice correspondence** (SCC) that assigns a set of social states for each possible profile. Any single-valued SCC will be a **social choice function** (SCF) and will be denoted by $f : \prod_{i=1}^{n} \Theta_i \to K$. A **social welfare function** (SWF) will be a real-valued function of the type: $W : K \times \prod_{i=1}^{n} \Theta_i \rightarrow E$, where $E$ is the real line. We will say that SCC $K^*$ is **generated** or represented by SWF $W$ iff $W(K^*(\theta), \theta) \geq W(k, \theta)\ \forall k \in K,\ \forall \theta \in \prod_{i=1}^{n} \Theta_i$, and $\overline{W}(K^*)$ will be the set of SWFs representing a given SCC $K^*$.[10].Given an economy $e$, suppose now that individual objectives given the type take the form of a **preference relation** on the set of alternatives, denoted by $R_i(\theta_i),\ \forall \theta_i \in \Theta_i$, that is, $R_i\ \forall i \in N$ is a mapping from the possible types to the set of all ordered pairs of alternatives: $R_i : \Theta_i \rightarrow\rightarrow K \times K$.[11] We say that the domain is **unrestricted** iff every complete weak pre-ordering is admissible as a preference relation. Assuming that each agent's type is his own private information and society - or the social planner - cannot directly observe the true individual types, the rule has to be based on the revealed types rather than on the true individual types. We are interested in SCFs such that each agent has no any incentive to lie about his true type in any case - whatever types the rest of agents report to the planner and whatever be the agent's true type -. We say that a SCF is **strategy-proof** iff

$$\forall i \in N,\ \forall \theta \in \prod_{i=1}^{n} \Theta_i,\ \forall \theta_i' \in \Theta_i,\ f(\theta) R_i(\theta_i) f(\theta_i', \theta_{-i}).$$

Gibbard [4] and Satterthwaite [14] proved that whenever the domain is unrestricted and $\#range(f) \geq 3$, *the only strategy-proof social choice functions are dictatorial, i.e.,* $\exists i \in N$ *such that* $\forall \theta \in \prod_{i=1}^{n} \Theta_i,\ f(\theta) \in \begin{array}{c} \arg\max\ R_i(\theta) \\ \text{s.t. } k \in range(f) \end{array}$.

2. *Roberts'* [13] *Theorem.*

Let us consider an economy $e$ such that $K$ is a finite set. We furthermore assume that the agents' objectives are not defined by the preference ordering on $K$ associated to each type, but by a real-valued payoff of the kind:

$$\forall i \in N, \quad P_i : K \times E \times \Theta_i \rightarrow \quad E$$

Where $P_i$ is quasi-linear with respect to the second argument, intended to represent some out-the-model way to compensate the agent, so that we can write agent $i's$ payoff function in the form: $P_i(k, q_i, \theta_i) = v_i(k, \theta_i) + q_i,\ \forall k \in$

[10]Notice that for every SCC $K^*$, the set $\overline{W}(K^*)$ is non-empty: the constant SWF trivially represents every SCC.

[11]We are implicitly assuming that the economy is such that allow for the definition of individual preferences.

$K$, $\forall \theta_i \in \Theta_i$, $\forall q_i \in E$, where for all $i$, $v_i : K \times \Theta_i \to E$ is a real-valued function called the **valuation function** admitting every possible cardinal utility scale on the set of (finite) alternatives. We say now that $f$ is strategy-proof iff there exist bounded **compensation functions** $q_i : \prod_{i=1}^n \Theta_i \to E$, $\forall i \in N$, such that[12] $\forall i \in N$, $\forall \widehat{\theta} \in \prod_{i=1}^n \Theta_i$, $\forall \theta_i \in \Theta_i$,

$$v_i(f(\theta_i, \widehat{\theta}_{-i}), \theta_i) + q_i(\theta_i, \widehat{\theta}_{-i}) \geq v_i(f(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i) + q_i(\widehat{\theta}_i, \widehat{\theta}_{-i}) \qquad (6)$$

Roberts proved that whenever $range(f) = K$, *the only strategy-proof social choice functions come from maximizing some weighted sum of the agents' $v_i's$*, i.e.,

$\exists a_1, ..., a_n \in E_+$, $\sum_{i=1}^n a_i = 1$, such that $\forall \theta \in \prod_{i=1}^n \Theta_i$,

$f(\theta) \in \underset{k \in K}{\arg\max} \sum_{i=1}^n a_i v_i(k, \theta_i) + F(k)$ where $F : K \to E$ is any bounded real-valued function.

3. *Groves-Clarke mechanisms.* (Groves [6], [7] and [8], Clarke [2], Green & Laffont [5]).

Let us consider an economy such that $K$ is now some compact set in a topological space. Let the real-valued bounded $v_i's$ in the above framework for all $\theta_i \in \Theta_i$ be the set of all bounded upper-semi-continuous or continuous functions. The list of compensation functions (**mechanism**) $\{q_1, ..., q_n\}$ *implements by revelation* the SCC $K^*$ iff for every selection $f$ from $K^*$, i.e., $f(\widehat{\theta}) \in K^*(\widehat{\theta})$ $\forall \widehat{\theta} \in \prod_{i=1}^n \Theta_i$, (1) holds.[13]

Green & Laffont [5] proved that the only mechanisms that implement the SCC $K^*(\theta) = \underset{s.t.\ k \in K}{\arg\max} \sum_{i=1}^n v_i(k, \theta_i)$ are the **Groves' mechanisms**, i.e., those mechanisms in which all the compensation functions take the form:

$$\forall i \in N,\ q_i(\widehat{\theta}) = \sum_{j \neq i} v_j(f(\widehat{\theta}), \widehat{\theta}_j) + h_i(\widehat{\theta}_{-i}),\ \forall f(\widehat{\theta}) \in K^*(\widehat{\theta}) \qquad (7)$$

where $h_i : \prod_{j \neq i} \Theta_j \to E$ is any real-valued function independent of $\widehat{\theta}_i$.

[12] There is no difference in the interpretation of Robert's definition of strategy-proofness from the definition above provided that we realise that there are now two ways in which the planner can affect the agents' final payoff (and hence their incentives): via the chosen alternative (affecting the $v_i's$) and by varying the compensations (that only depend on the revealed profile), that is the direct way of changing the final payoffs.

[13] Again, there is no difference between the notions of strategy-proofness in Roberts' setup and the notion of mechanisms implementing by revelation a SCF.

Later research (Green & Laffont [5], Walker [15] and Hurwicz & Walker [11]) proved that any such mechanism cannot generically *balance the budget*, that is, $\forall \{q_1, ..., q_n\}$ implementing by revelation the SCC above, $\forall \overline{k} \in E$, $\sum_{i=1}^{n} q_i(\widehat{\theta}) \neq \overline{k} \; \forall \widehat{\theta} \in \prod_{i=1}^{n} \Theta_i$.

The three implementation scenarios described above share some common properties:

(i). The domains of private characteristics are quite *large* in each model.

(ii). The incentive compatibility requirement is actually the same in all cases and amounts to the existence of truth-revealing dominant strategies.

(iii). The implementable social choice rules or functions in the three cases require that individuals' objectives are made somehow *similar* to social objectives. Gibbard does not allow for extra-model compensations that can affect the agents' objectives but not the planner's utility -any SWF representing the SCC-, so he obtains that implementability should make the social objectives be identified with those of some fixed agent -the dictator-. Roberts, in his turn, allows for *monetary* or quasi-linear compensations in the mechanism design, but the only social choice functions that are implementable come from the maximization of some linear combination of the individuals' valuation functions -some *quasi-linear* or utilitarian social objectives-. Green & Laffont actually work in the same framework of Roberts, but they are interested in implementing the utilitarian social welfare function -the sum of all the agents' valuation functions- and using quasi-linear individual objectives. Notice that the Groves' compensation functions -the only mechanisms working in that domain- somehow replicate the social objectives in the sense that if the planner delegate the outcome selection in any agent, he would actually pick up the same social alternatives that would be chosen by the planner himself.

The analogies pointed out above give rise to a conjecture that might be formalized as follows:

**Conjecture 48** *Truthful implementation in* large *domains of SCC $K^*$ requires that $\forall i \in N$, for all* admissible *functions $P_i : E^2 \to E$, the following holds: $\forall \theta_i, \widehat{\theta}_i \in \Theta_i, \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$,*

$$P_i(v_i(f(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = W(f(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i}).$$

*for any $W \in \overline{W}(K^*)$.*

It is not difficult to check that the three results summarized above are particular cases of the above statement -or can be written in these terms-.

We will then raise the following questions: Is that a common feature of all truthful implementation problems in large domains?. How far can this conjecture be extended? Is it valid for every compensation scheme, even for those not restricted to quasi-linear compensations?

We find that there exists a requirement on the social choice rule called **individual decisiveness**, under which truthful implementation demands social and individual objectives to coincide in the above sense. This property assumes a strong responsiveness of the social choices to changes on the individual types. Examples will be provided later, but let us point out now that individual decisiveness holds for the usual social choice rules when allowing sufficiently rich domains. In particular, it holds and plays a crucial role in environments admitting Groves' type mechanisms.

The main result obtained in this chapter shows the strong connection between the specific agents' payoff function structure and the individually decisive social choice rules that can be truthfully implemented in dominant strategies. In summary, the compensations allowed in any mechanism should be such as to allow that the payoff function structure replicates some social welfare function representing that social choice rule. In other words, we should give to the agents exactly the same incentive scheme that the one implied in the social welfare function; the objective function of agents and that of society should somehow coincide.

The remainder of the chapter proceeds as follows: first we introduce the model with the definitions, the main result is established in *Section 3* and *Section 4* is devoted to applications in different environments. Conclusions follow.

## 4.2 The model

We shall propose a general model, including the setups for the examples of the preceding section as particular cases. *Example 1* (the Gibbard-Satterthwaite Theorem) occurs in a setup which is traditional in social choice theory: alternatives are defined as those objects over which agents are assumed to have preferences. *Examples 2* and 3 (Robert's and Groves-Clarke's) refer to setups where social states are described through two sets of variables: the levels of variables in the first set are interpreted as the result of a *public decision*; the levels of others are interpreted as transfers of goods that may be interpreted as compensations to agents. Individual payoffs depend on the overall levels of both types of variables. If we wanted to keep the conventions of social choice we should reserve the term alternative to denote these combinations of levels, since it is on them that agents have preferences. But it is more useful to keep the distinction found in the other two models, and eventually to generalize it. Hence, we distinguish between those parts of a decision which involve a public decision ($K$) and those ($q$) which can be interpreted as compensations for the agents. The agent's valuations of the public decisions will be assumed to be well defined as functions of their types, and the overall preferences of agents over public decisions and compensation levels will be assumed to take a not necessarily additive form. This will make our setup more general than any of those mentioned above.

Let us consider **any economy** $e$. Given $e$, we define the following:

**Definition 49** *A **compensation mechanism** $\{P, q\}$ is the tuple defined by the following sets:*

*(i) $P = \{P_i, \forall i \in N\}$ is the set of **payoff functions**[14], where $\forall i \in N,\ P_i : E^2 \to E$ is a continuous upper-bounded real-valued function* monotonic *in both arguments -and strictly monotonic in the second-, i.e.,*

$\forall x, x', y, y' \in E,\ \ y > y' \Rightarrow P_i(x, y) > P_i(x, y')$ *and* $x > x' \Rightarrow P_i(x, y) \geq P_i(x', y)$.

*(ii) $q = \{q_i, \forall i \in N\}$ is the list of **compensation functions**: $\forall i \in N,\ q_i : \prod_{i=1}^{n} \Theta_i \to E$, upper-bounded real-valued functions that serve the planner to distribute utility among the agents based in the information contained in the strategies.*

Given a compensation mechanism, the **final payoff** that any individual gets from any strategy vector is always given by the following expression:

$$\forall i \in N,\ \pi_i(\widehat{\theta}, \theta_i) = P_i(v_i(g(\widehat{\theta}), \theta_i), q_i(\widehat{\theta})).$$

Therefore the functional form of the final payoff is partially given by the mechanism and does not necessarily coincide with the valuation function, except in the limit case of a mechanism such that $\forall i \in N,\ \forall x, y_i \in E,\ P_i(x, y_i) =$

[14] These functions stand for the payoffs that the agents get, given their types -be it a production function, a utility function or a general arbitrary type of agent-, and some individual real argument which can be used within the mechanism in order to compensate the agent.

$x$, which is not strictly monotonic in the second argument. The specific payoff function structure will allow us to classify every compensation mechanism. In order to illustrate this point, it will be useful to think about a production economy: $e$ is such that $N$ is a set of firms, divisions within a firm or productive agents that produce a single homogeneous good, $K$ represents either feasible levels of a public input used by the agents or distributions of a private input; let us assume that the set of types determines the feasible technologies available such that the valuation function will be a production function. Consider the following examples of particular payoff functions: $\forall x, y_i \in E,\ P_i(x, y_i) = x,\ P_i(x, y_i) = x + y_i,\ P_i(x, y_i) = xy_i$, and $P_i(x, y_i) = y_i$. Those payoff functions imply final payoffs of the form: $\pi_i = v_i(g(\widehat{\theta}), \theta_i),\ \pi_i = v_i(g(\widehat{\theta}), \theta_i) + q_i(\widehat{\theta}),\ \pi_i = v_i(g(\widehat{\theta}), \theta_i)q_i(\widehat{\theta}),\ \pi_i = q_i(\widehat{\theta})$ respectively. The first one represents the impossibility of compensating agents. Agent $i's$ objective function is fully determined given by his own private characteristic -or his produced output for given levels of the input-, so it coincides with the usual implementation framework, where every possible compensation is modelled *inside* the set of feasible alternatives. We will refer to this case as the *compensation free payoff functions.* In the second example above, the productive agent sells his output at some given (unitary) price and gets the profit, but the planner or principal can only set some tax or subsidy to provide an incentive for truthful behavior. This will be called the *compensation by transfers* case. The third example assumes the ability of the planner to set the final price of the produced good according to some pre-specified rule -*compensation by prices*-. Finally, in the last example the agent has no property rights on the good produced, and he only receives a wage that can depend on the information reported by all the agents. This will be the *full compensation* case. Notice that all the above examples allow for different compensation or surplus-sharing schemes and the possibility of one or another may be made discretional to the planner in some contexts or given by nature in others. Furthermore, the monotonicity property imposed on the payoff functions establishes a specific restriction on the functional form of compensations, so that for any given allocation, the higher the compensation, the higher the final payoff. A non-monotonic payoff function might be, for example, the following one: $P_i(x, y) = xy_i^2$. This condition does not seem to be too restrictive, since the specific nature of the compensation requires a clear guide to reward the agents.

Finally, notice that the compensation functions can also be viewed as part of the *real alternatives*, and we could define an extended set of alternatives as: $K' = K \times E^n$, where compensations for all the agents are included within the set of alternatives. Once the payoff function structure is fixed, it amounts exactly to the usual framework of compensation free payoff functions with the agents' valuations re-defined to be: $\forall i \in N,\ v_i'(k', \theta_i) = P_i(v_i(k, \theta_i), y_i)$ and $k' = (k, y_1, .., y_n)$, so the approach followed in most of the implementation literature -see, for example, Green and Laffont [5]- is a particular case of our compensation mechanisms too.

Now we will define the notion of truthful implementation that we will use, together with some additional definitions related to SCCs and mechanisms.

**Definition 50** *Given an economy $e$ and a SCC $K^*$, we say that a compensation mechanism $\{P, q\}$ is an* ***incentive compatible mechanism for*** *$i$ if the following condition holds for $i : \forall \theta_i, \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$,*

$$P_i(v_i(f(\theta_i, \widehat{\theta}_{-i}), \theta_i), q_i(\theta_i, \widehat{\theta}_{-i})) \geq P_i(v_i(f(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})), \quad (8)$$

*and this for any selection $f(\widehat{\theta})$ from $K^*(\widehat{\theta})$.*

A compensation mechanism which is incentive compatible for all $i \in N$ is said to be *incentive compatible*. If an incentive compatible mechanism exists for some SCC we say that the mechanism *implements by revelation* that SCC.

The last definitions are natural generalizations of those in Green and Laffont [5].

**Definition 51** *A SCC $K^* : \prod_{i \in N} \Theta_i \to K.$ is called* ***individually decisive for*** *$i$ if*

$$\forall \theta_{-i} \in \Theta_{-i},\ \ \forall k \in K,\ \exists \widehat{\theta}_i \in \Theta_i\ \ni\ k \in K^*(\widehat{\theta}_i, \theta_{-i}).$$

*A SCC that is individually decisive for all $i$ is individually decisive and any SWF which represents some individually decisive SCC will be individually decisive too.*

This property means that individual $i$ can force any alternative under some circumstances to be in the choice set by declaring an appropriate characteristic for any others' types.

It may be useful to illustrate the above definitions with some examples of both individually decisive and non-decisive SCCs.

**Example 52** *Let the economy e be such that K is a compact set defined in a topological space and* $\forall i \in N$, $\Theta_i$ *includes all the bounded, upper-semi-continuous functions on K. Let us consider the utilitarian welfare function:* $W = \sum_{i=1}^{n} v_i(k, \theta_i)$. *The SCC that maximizes W on K, i.e.,* $K^*(\theta_1, ..., \theta_n) = \underset{k \in K}{\arg\max} \; \sum_{i=1}^{n} v_i(k, \theta_i)$ *, can be proved (see* Proposition 66 *later) to be individually decisive.*

**Example 53** *The Pareto SCC is individually decisive when defined on a large set of rich domains;* $\forall \theta \in \prod_{i \in N} \Theta_i$,

$$PO(\theta) = \left\{ \begin{array}{c} k \in K \text{ such that } \nexists \overline{k} \in K \text{ such that } \forall i \in N,\ v_i(\overline{k}, \theta_i) \geq v_i(k, \theta_i) \\ \text{and } v_i(\overline{k}, \theta_i) > v_i(k, \theta_i) \text{ for some } i \in N \end{array} \right\}.$$

*Consider the unrestricted domain on* $K : \forall i \in N$, $\Theta_i$ *is such that* $\forall k, l \in K$ $(l \neq k)$, $\exists \widehat{\theta}_i \in \Theta_i$ *such that* $v_i(k, \widehat{\theta}_i) > v_i(l, \widehat{\theta}_i) \Rightarrow \forall k \in K$, $\exists \widehat{\theta}_i \in \Theta_i$ *such that:* $k \in PO(\widehat{\theta}_i, \theta_{-i}) = K^*(\widehat{\theta}_i, \theta_{-i})$.

**Example 54** *Consider now the Pareto SCC in an economy such that* $N = \{1, 2\}$ *and the domain of all continuous, strictly monotonic and convex preference orderings over the 2-good commodity space* $E_+^2$ *when the set K is the set of all feasible allocations of the 2 goods available in fixed finite amounts between the two agents (the Edgeworth Box). Take any* $\theta_2 \in \Theta_2$ *and any feasible allocation* $z \in K$. *Construct the agent 2's upper contour set on* $z : C_2(z, \theta_2) = \{y \in K \text{ s.t. } v_2(y, \theta_2) \geq v_2(z, \theta_2)\}$, *and find, for example, the value "a"* $\in E_+$ *s.t.*

$$z \in \underset{x}{\arg\max} \; a x_1^1 + x_2^1 \quad \text{s.t.} \quad x \in C_2(z, \theta_2),\ x \in K$$

*. This value will always exist because of the convexity and monotonicity assumptions on* $\Theta_1$ *and* $\Theta_2$. *Now, set* $\widehat{\theta}_1 = a$ *and define* $v_1(k, \widehat{\theta}_1) = a x_1^1 + x_2^1 \ \forall x_1^1, x_2^1$. *This is a convex, strictly monotonic and continuous function, so* $\widehat{\theta}_1 \in \Theta_1$, *and by construction,* $z \in PO(\widehat{\theta}_1, \theta_2) = K^*(\widehat{\theta}_1, \theta_2)$.

**Example 55** *In an economy like that in* Example 54*, The Walrasian correspondence with respect to some vector of initial endowments can be proved to be an individually decisive SCC by using a similar argument.*

**Example 56** *Let us assume an economy such that* $n \geq 3$ *is odd, K is a closed interval of the real line* $K = [0, 1] \subset E$ *and the agents' types are such* *that the agents may have every continuous single-peaked valuation function, i.e.,* $\forall i \in N,\ \Theta_i$ *is such that* $\forall \theta_i \in \Theta_i,\ \exists \overline{k}(\theta_i) \in K \quad s.t.\ v_i(\overline{k}(\theta_i), \theta_i) > v_i(k, \theta_i)\ \forall k \in K \backslash \overline{k}(\theta_i)\ \&\ \forall k^{'}, k^{''} \in K\ (k^{'} > k^{''})$,

$\overline{k}(\theta_i) \geq k^{'} \Rightarrow v_i(k^{'}, \theta_i) \geq v_i(k^{''}, \theta_i)\ \&\ \ \overline{k}(\theta_i) \leq k^{''} \Rightarrow v_i(k^{'}, \theta_i) \leq v_i(k^{''}, \theta_i)$. *Consider the following SCC (the* median voter *SCC):* $\forall \theta \in \prod_{i=1}^{n} \Theta_i,\ K^*(\theta) = med_{i \in N} \left\{ \overline{k}(\theta_i) \right\}$, *where the function "med" stands for the median of the revealed* peaks *of the agents, i.e., the peak such that leaves the same number of other peaks on the right and on the left. This is a well-known selection of the Pareto correspondence in this economy -see, for example, Moulin* [12]*-. It is easy to see that there exist situations where some individuals cannot individually change the decision, for example, individual* $i$ *cannot even affect the decision for* $\theta_{-i} \in \Theta_{-i}$ *such that* $\overline{k}(\theta_j) = \overline{k}(\theta_h)\ \forall j, h \neq i$, *so* $K^*$ *is not individually decisive.*

**Example 57** *Let us consider an economy* $e$ *such that* $K = [0,1] \subset E_+$ *and* $\forall i \in N,\ \Theta_i = E_+$, *and* $v_i(k, \theta_i) = a_i k - \frac{1}{2}k^2\ \forall k \in K, \forall a_i \in \Theta_i$. *The utilitarian SWF in* Example 52 *above is individually non-decisive, but it will be individually decisive if the economy is enlarged to allow for types such that* $\Theta_i = E$ *(actually, allowing for convex valuation functions). Notice that in this cases,* $K^*(a_1, ..., a_n) = \frac{1}{n}\sum_{i=1}^{n} a_i$.

We can now address the main question we face: what kind of compensation mechanisms, if they exist, should we use in order to implement by revelation any individually decisive SCC?. We will provide a complete answer to this question.

## 4.3 Main result

The results presented below are all of them valid for every given economy $e = \langle N, K, \Theta_i\ \forall i \in N \rangle$.

**Theorem 58** *The only incentive compatible compensation mechanisms* $\{P, q\}$ *that implement by revelation an individually decisive SCC* $K^*$ *are such that:* $\forall i \in N,\ \forall \theta_i, \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$,

$$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = W(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i}),$$

*for every selection* $g$ *from* $K^*$.

In order to prove the theorem, we will make use of the following intermediate results:

**Lemma 59** *Let $\{P,q\}$ be an incentive compatible for $i$ compensation mechanism implementing the SCC $K^*$, and take two -possibly the same- selections from $K^*$ : $g$ and $\widehat{g}$. Then, $\forall\theta_i,\theta_i'\in\Theta_i,\ \forall\theta_{-i}\in\prod_{j\neq i}\Theta_j$, s.t. $g(\theta_i,\theta_{-i})=\widehat{g}(\theta_i',\theta_{-i})$, it holds that $q_i(\theta_i,\theta_{-i})=q_i(\theta_i',\theta_{-i})$.*

**Proof.** Suppose the contrary, i.e., $\exists\theta_{-i}\in\prod_{j\neq i}\Theta_j,\ \exists\theta_i,\theta_i'\in\Theta_i$ such that both yield the same outcome with both selections: $g(\theta_i,\theta_{-i})=\widehat{g}(\theta_i',\theta_{-i})=\overline{k}$, and one of them leads to a bigger compensation: $\theta_i$ w.l.g., then: $q_i(\theta_i,\theta_{-i})>q_i(\theta_i',\theta_{-i})$. Then, take type $\theta_i'$, and consider the payoffs:

$P_i(v_i(g(\theta_i,\theta_{-i}),\theta_i'),q_i(\theta_i,\theta_{-i}))$ and $P_i(v_i(\widehat{g}(\theta_i',\theta_{-i}),\theta_i'),q_i(\theta_i',\theta_{-i}))$. By hypothesis, the outcome will be the same: $g(\theta_i,\theta_{-i})=\widehat{g}(\theta_i',\theta_{-i})=\overline{k}\Rightarrow$

$\Rightarrow v_i(g(\theta_i,\theta_{-i}),\theta_i')=v_i(\widehat{g}(\theta_i',\theta_{-i}),\theta_i')$. Then, by monotonicity of the payoff functions structure, we should have:

$P_i(v_i(g(\theta_i,\theta_{-i}),\theta_i'),q_i(\theta_i,\theta_{-i}))>P_i(v_i(\widehat{g}(\theta_i',\theta_{-i}),\theta_i'),q_i(\theta_i',\theta_{-i})).\Rightarrow$

For $\theta_i'$, $\exists\theta_i\in\Theta_i$, $\exists\theta_{-i}\in\prod_{j\neq i}\Theta_j$, $\exists\overline{g}(\theta_i,\theta_{-i})\in K^*(\theta_i,\theta_{-i})\ \forall\theta_i\in\Theta_i$, $\forall\theta_{-i}\in\prod_{j\neq i}\Theta_j$, - a selection from $K^*$- defined as:

$$\overline{g}(\widehat{\theta}_i,\theta_{-i})=\begin{cases} g(\widehat{\theta}_i,\theta_{-i}) & iff\ \widehat{\theta}_i\neq\theta_i' \\ \widehat{g}(\theta_i',\theta_{-i}) & iff\ \widehat{\theta}_i=\theta_i'\end{cases}$$

such that by declaring the last one we are better than reporting the true characteristic, i.e.,

$P_i(v_i(\overline{g}(\theta_i,\theta_{-i}),\theta_i'),q_i(\theta_i,\theta_{-i}))>P_i(v_i(\overline{g}(\theta_i',\theta_{-i}),\theta_i'),q_i(\theta_i',\theta_{-i}))$, so the mechanism cannot be incentive compatible for $i$.[15] ■

**Lemma 60** *The only incentive compatible for $i$ compensation mechanisms $\{P,q\}$ that implement by revelation any individually decisive for $i$ SCC $K^*$ are such that: $\forall\theta_i,\widehat{\theta}_i\in\Theta_i,\ \forall\widehat{\theta}_{-i}\in\prod_{j\neq i}\Theta_j$,*

$$P_i(v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\theta_i),q_i(\widehat{\theta}_i,\widehat{\theta}_{-i}))=W(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\theta_i,\widehat{\theta}_{-i}),$$

*for any selection $g$ from $K^*$.*

**Proof. Necessity** $\Rightarrow$) We suppose that $\{P,q\}$ is an incentive compatible for $i$ compensation mechanism implementing by revelation some individually

[15] This lemma is a generalization of part of Green and Laffont's [5] Theorem 3.

decisive for $i$ SCC $K^*$. Since $K^*$ is individually decisive for $i$, it is true that $\forall \theta_{-i} \in \Theta_{-i},\ \forall k \in K,\ \exists \widehat{\theta}_i(k, \theta_{-i}) \in \Theta_i \ \ni \ k \in K^*(\widehat{\theta}_i, \theta_{-i})$. Hence, the mapping $\widehat{\theta}_i : K \times \prod_{j \neq i} \Theta_j \to\to \Theta_i$ is well-defined and for any selection $\overline{\theta}_i$ from that mapping, it holds that $\forall k \in K, \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j,\ k \in K^*(\overline{\theta}_i(k, \theta_{-i}), \theta_{-i})$. Now, let us define the following mapping for individual $i \in N$: $\widehat{q}_i : K \times \prod_{j \neq i} \Theta_j \to\to E$. defined as follows: $\forall k \in K,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$,

$\widehat{q}_i(k, \theta_{-i}) = \left\{ q_i(\overline{\theta}_i(k, \theta_{-i}), \theta_{-i}), \text{ for any selection } \overline{\theta}_i \text{ from } \widehat{\theta}_i \right\}$. This mapping is well defined and has the following properties:

i) $range(\widehat{q}_i) = range(q_i)$.

ii) $dom(\widehat{q}_i) = K \times \prod_{j \neq i} \Theta_j$.

iii) $\widehat{q}_i$ is a real-valued *function*.

iv) $\forall \theta_i \in \Theta_i,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j,\ \ \widehat{q}_i(g(\theta_i, \theta_{-i}), \theta_{-i}) = q_i(\theta_i, \theta_{-i})$, for any selection $g$ from $K^*$.

Property (i) is obvious by definition: for any $\theta_i \in \Theta_i$, $\exists k = g(\theta_i, \theta_{-i})$, so $q_i(\theta_i, \theta_{-i}) \in \widehat{q}_i(k, \theta_{-i})$. (ii) holds because $\widehat{\theta}_i$ is a well-defined mapping from $K \times \prod_{j \neq i} \Theta_j$. Property (iii) holds because we are in the conditions of applying *Lemma 59*: Since $\{P, q\}$ is an incentive compatible for $i$ compensation mechanism implementing the SCC $K^*$ by hypothesis, $\forall \theta_{-i} \in \prod_{j \neq i} \Theta_j,\ \forall \theta_i, \theta_i'$ $\in \Theta_i$ s.t. $g(\theta_i, \theta_{-i}) = \widehat{g}(\theta_i', \theta_{-i}) = k$ for two -possibly the same- selections from $K^* \Rightarrow q_i(\theta_i, \theta_{-i}) = q_i(\theta_i', \theta_{-i})$. Thus, $\forall \theta_{-i} \in \prod_{j \neq i} \Theta_j,\ \forall k \in K$, $q_i(\overline{\theta}_i(k, \theta_{-i}), \theta_{-i}) = q_i(\widetilde{\theta}_i(k, \theta_{-i}), \theta_{-i})$ for any two arbitrary elections $\overline{\theta}_i$ and $\widetilde{\theta}_i$ from $\widehat{\theta}_i$ and hence a single real number is associated to each $k \in K$ in the function $\widehat{q}_i(k, \theta_{-i})$.

Finally, property (iv) is straightforward by definition and (iii).

Now, we know by assumption that $\{P, q\}$ is an incentive compatible compensation mechanism for $i$, so that it holds that:

$P_i(v_i(g(\theta_i, \theta_{-i}), \theta_i), q_i(\theta_i, \theta_{-i})) \geq P_i(v_i(g(\theta_i', \theta_{-i}), \theta_i), q_i(\theta_i', \theta_{-i}))$

$\forall \theta_i, \theta_i' \in \Theta_i,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$ and this for any selection $g(\theta_i, \theta_{-i}) \in K^*(\theta_i, \theta_{-i})$.

Now, using iv), we can state the following:

$P_i(v_i(g(\theta_i, \theta_{-i}), \theta_i), \widehat{q}_i(g(\theta_i, \theta_{-i}), \theta_{-i})) \geq$

$\geq P_i(v_i(g(\theta_i', \theta_{-i}), \theta_i), \widehat{q}_i(g(\theta_i', \theta_{-i}), \theta_{-i}))\ \ \forall \theta_i, \theta_i' \in \Theta_i,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$ and for any selection $g(\theta_i, \theta_{-i}) \in K^*(\theta_i, \theta_{-i})$.

Then, since $K^*$ is individually decisive by hypothesis, we can choose the following selection $\widetilde{g}(\theta_i', \theta_{-i})$ from $K^*$: Given any $\theta_{-i} \in \prod_{j \neq i} \Theta_j$ and any $\theta_i \in \Theta_i,\ \forall \theta_i' \in \Theta_i$,

$$\widetilde{g}(\theta'_i,\theta_{-i}) = \begin{cases} k & iff\ \theta'_i = \overline{\theta}_i(k,\theta_{-i})\ \&\ \theta'_i \neq \theta_i \\ g(\theta'_i,\theta_{-i}) & otherwise \end{cases}, \quad (2).$$

where $\overline{\theta}_i(k)$ is any selection from $\widehat{\theta}_i$ and $g(\theta'_i,\theta_{-i})$ is any arbitrary selection from $K^*$. It holds for this selection that: $\forall\theta_i,\theta'_i \in \Theta_i$,

$P_i(v_i(\widetilde{g}(\theta_i,\theta_{-i}),\theta_i),\widehat{q}_i(\widetilde{g}(\theta_i,\theta_{-i}),\theta_{-i})) \geq$

$P_i(v_i(\widetilde{g}(\theta'_i,\theta_{-i}),\theta_i),\widehat{q}_i(\widetilde{g}(\theta'_i,\theta_{-i}),\theta_{-i}))$. (3). But this is true for all $\theta'_i \in \Theta_i$, which only affects the right hand side of the above inequality and, by definition of $\widetilde{g}$, $\{\widetilde{g}(\theta'_i,\theta_{-i}),\ \forall\theta'_i \in \Theta_i\} = K$, so for each $\theta_i \in \Theta_i$,we can write (3) in the following way: given any $\theta_i$ and $\theta_{-i}$, we can construct a selection $\widetilde{g}$ defined above and obtain:

$P_i(v_i(g(\theta_i,\theta_{-i}),\theta_i),\widehat{q}_i(g(\theta_i,\theta_{-i}),\theta_{-i})) \geq P_i(v_i(k,\theta_i),\widehat{q}_i(k,\theta_{-i}))\ \ \forall k \in K$. (4).

But $g(\theta_i,\theta_{-i})$ for each $\theta_i \in \Theta_i$ is selected arbitrary from $K^*(\theta_i,\theta_{-i})$, while the right hand side of the above inequality is the same for each selection given $\theta_i$ and $\theta_{-i}$, so statement (4) holds for every selection from $K^*(\theta_i,\theta_{-i})$. Abusing notation, we can write (4) as follows:

$\forall\theta_i \in \Theta_i$,

$P_i(v_i(K^*(\theta_i,\theta_{-i}),\theta_i),\widehat{q}_i(K^*(\theta_i,\theta_{-i}),\theta_{-i})) \geq P_i(v_i(k,\theta_i),\widehat{q}_i(k,\theta_{-i}))\ \ \forall k \in K$. (5).

Now, let us consider the following composed function $\widehat{P}_i : K \times \prod_{i=1}^n \Theta_i \to E$ defined as: $\forall k \in K,\ \forall\theta \in \prod_{j=1}^n \Theta_j,\ \ \widehat{P}_i(k,\theta_i,\theta_{-i}) = P_i(v_i(k,\theta_i),\widehat{q}_i(k,\theta_{-i}))$, (6), which is well-defined when $K^*$ is individually decisive. Notice that (5) can be written -slightly abusing notation again- as:

$\forall k \in K,\ \forall\theta_i \in \Theta_i,\ \forall\theta_{-i} \in \prod_{j\neq i}\Theta_j,\ \widehat{P}_i(K^*(\theta_i,\theta_{-i}),\theta_i,\theta_{-i}) \geq \widehat{P}_i(k,\theta_i,\theta_{-i})$.

But this last expression is the definition of some SWF representing SCC $K^*$. In other words, let us suppose that $\widehat{P}_i \notin \overline{W}(K^*)$. This can only be true when $\exists\theta_i \in \Theta_i,\ \exists\widetilde{\theta}_{-i} \in \prod_{j\neq i}\Theta_j,\ \exists\widetilde{k} \in K$ such that $\widehat{P}_i(\widetilde{k},\theta_i,\widetilde{\theta}_{-i}) > \widehat{P}_i(K^*(\theta_i,\widetilde{\theta}_{-i}),\theta_i,\widetilde{\theta}_{-i})$ (7). But since $K^*$ is individually decisive for $i$, for $\widetilde{\theta}_{-i}$ there exist $\exists\widetilde{\theta}_i \in \Theta_i$ such that $\widetilde{k} \in K^*(\widetilde{\theta}_i,\widetilde{\theta}_{-i})$. Substituting this into (7), we have found a selection of $K^*$ such that, slightly abusing notation again, $\exists\widetilde{\theta}_i \in \Theta_i$ $(\widetilde{\theta}_i \neq \theta_i)$, $\exists\theta_i \in \Theta_i$, $\exists\widetilde{\theta}_{-i} \in \prod_{j\neq i}\Theta_j$, such that

$\widehat{P}_i(K^*(\widetilde{\theta}_i,\widetilde{\theta}_{-i}),\theta_i,\widetilde{\theta}_{-i}) > \widehat{P}_i(K^*(\theta_i,\widetilde{\theta}_{-i}),\theta_i,\widetilde{\theta}_{-i})$, and, which by (6) can be written again as: $P_i(v_i(K^*(\widetilde{\theta}_i,\widetilde{\theta}_{-i}),\theta_i),\widetilde{\theta}_{-i}) > P_i(v_i(K^*(\theta_i,\widetilde{\theta}_{-i}),\theta_i),\widetilde{\theta}_{-i})$, and this clearly contradicts mechanism $\{P,q\}$ to be incentive compatible for $i$. Hence, it has to be that $\widehat{P}_i \in \overline{W}(K^*)$ and the existence is proved.

**Sufficiency** $\Leftarrow$) Now we have to prove that every mechanism such that $\forall\theta_i,\widehat{\theta}_i \in \Theta_i,\ \forall\widehat{\theta}_{-i} \in \prod_{j\neq i}\Theta_j$, can be written as $P_i(v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\theta_i),q_i(\widehat{\theta}_i,\widehat{\theta}_{-i})) =$

$W(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i})$, (1) is an incentive compatible for $i$ compensation mechanism. Suppose, on the contrary, that $\exists g \in K^*$, $\exists\widetilde{\theta}_i \in \Theta_i$ , $\exists\widetilde{\theta}_{-i} \in \prod_{j\neq i} \Theta_j$ and $\exists\theta'_i \in \Theta_i$ such that $W(g(\theta'_i, \widetilde{\theta}_{-i}), \widetilde{\theta}_i, \widetilde{\theta}_{-i}) > W(g(\widetilde{\theta}_i, \widetilde{\theta}_{-i}), \widetilde{\theta}_i, \widetilde{\theta}_{-i})$ (3). But, since $W$ represents $K^*$, it must be that $\forall\theta_i \in \Theta_i$, $\forall\theta_{-i} \in \prod_{j\neq i} \Theta_j$, $\forall k \in K$, $W(g(\theta_i, \theta_{-i}), \theta_i, \theta_{-i}) \geq W(k, \theta_i, \theta_{-i})$, and, in particular,

$W(g(\widetilde{\theta}_i, \widetilde{\theta}_{-i}), \widetilde{\theta}_i, \widetilde{\theta}_{-i}) \geq W(k, \widetilde{\theta}_i, \widetilde{\theta}_{-i})$ (4). take any $k = g(\theta'_i, \widetilde{\theta}_{-i}) \in K$, and (3) and (4) imply:

$W(g(\theta'_i, \widetilde{\theta}_{-i}), \widetilde{\theta}_i, \widetilde{\theta}_{-i}) > W(g(\widetilde{\theta}_i, \widetilde{\theta}_{-i}), \widetilde{\theta}_i, \widetilde{\theta}_{-i}) \geq W(g(\theta'_i, \widetilde{\theta}_{-i}), \widetilde{\theta}_i, \widetilde{\theta}_{-i})$, a contradiction, so $\{P, q\}$ is an incentive compatible compensation mechanism for $i$ and the lemma is proved. ■

**Proof of Theorem 58**:

Using *Lemma 60* and applying it for all $i$, it holds trivially.

The implications of *Theorem 58* are wide: It shows, for example, that when the social objectives are flexible enough, like the set of all continuous preferences on some compact set of alternatives, and we are trying to implement selections of the Pareto-optimal correspondence, which is clearly an individually decisive SCC, we must make coincide social interest with every individual's to achieve a positive result. Notice that it is a generalization of the well-known Groves' mechanisms: Green & Laffont's [5] result can be seen as a corollary of this one, and, moreover, it shows that the only restriction on preferences that allow for efficient and strong incentive compatible implementation are the quasi-linear domain.

## 4.4 Applications

In what follows, we will be concerned with different applications of *Theorem 58* in different contexts that fit their assumptions. We will show that a lot of interesting economic environments match our general model and our result will be very useful to characterize the mechanisms and social choice rules that allow for implementability. We classify the different applications by the compensation mechanism allowed.

### 4.4.1 Compensation-free schemes

**Corollary 61** *The only individually decisive for i SCCs that can be implemented by means of compensation mechanism when the payoff function structure is compensation-free is the dictatorial one.*

**Proof.** The compensation-free payoff function structure is not actually monotonic, but we do not need monotonicity to hold in this special case. Following the reasoning in *Theorem 58*, it is easy to see that *Lemma 59* is not necessary to prove that *Theorem 58* holds in this particular context, so the only compensation mechanism that implements by revelation any individually decisive for $i$ SCC is such that the following holds: $\forall\theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall\widehat{\theta}_{-i} \in \prod_{j\neq i}\Theta_j$,

$$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i) = W(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i})$$

for any $W \in \overline{W}(K^*)$. So we can only implement the SWF that represents the characteristic of individual $i$, and whatever any other individual reports, the SCC will be: $K^*(\theta_i, \theta_{-i}) = \underset{k}{argmax} \;\; v_i(k, \theta_i) \;\; \forall\theta_{-i} \in \Theta_{-i}$, i.e., individual $i$ is a dictator. ■

*Corollary 61* is a stronger version of Gibbard-Satterthwaite famous Theorem since we are imposing additional restrictions on the SCC to get the result -the SCC or SCF should be individually decisive for $i$-, which is a much stronger assumption that Gibbard's condition -the range of the function contains at least three elements-. It is related too with Barberà-Peleg's [1] version when considering continuous preferences and with *Theorem 3.2* in Roberts [13], but the case of compensation-free payoff functions becomes trivial in our framework[16] and the reason for considering this case here is to compare the gains from the possibilities of different compensation schemes below with the radical result of not allowing any kind of compensation.

**Corollary 62** *Consider any economy where* $\Theta_i$, $\forall i \in N$ *are such that* $\exists i, j \in N$, $\exists\theta_i \in \Theta_i, \exists\theta_j \in \Theta_j$ *such that* $\underset{k}{argmax} \;\; v_i(k, \theta_i) \neq \underset{k}{argmax} \;\; v_j(k, \theta_j)$ . *Then, there does not exist any individually decisive SCC that can be implemented by revelation by a compensation-free mechanism.*

**Proof.** Trivial: There cannot be more than one dictator when extending *Corollary 61* to more agents. ■

The compensation-free scheme is also interesting because usual mechanisms in the literature do not allow for compensations that are not included in

[16] Note that Barberá & Peleg's [1] proof of Gibbard-Satterthwaite Theorem consists of proving that strategy-proofness alone in an unrestricted domain of preferences makes the SCF either individually decisive for each agent or invariant to changes of his type.

the set of feasible alternatives. If this is the case and the planner cares about the whole payoff that the agents receive, our approach remains valid if we restrict the analysis to the compensation-free scheme. Notice that *Theorem 58* hold for the compensation-free case, and this will allow us to deal with the problem of balance -omitted until now- and provide a partial answer to an important question posed by Hurwicz &Walker [11]: Does an incentive compatible mechanism implementing the Pareto-optimal SCC exist for the mixed economy -public and private goods- when we drop the quasi-linear payoffs assumption?. In their own words, "*First, there is no reason to believe that* [their result] *depends upon the quasi-linearity of the individual's preferences; however, it is not clear how to obtain the result without the quasi-linearity assumption.*" The former authors could not directly face that problem because they used a characterization theorem due to Holmstrom which only holds for those payoffs, but we will show in what follows that our previous results are actually a powerful tool to deal with the general problem.

Let us consider a particular and simplified 2-agents mixed economy. There are two goods: one public and other private. Let $Y = \left[\underline{y}, \overline{y}\right]$ $, 0 < \underline{y} < \overline{y} < \infty$, be some compact interval on the real line representing the quantity provided of the public good and let $X = Y \times X_1 \times X_2$ be the consumption space in the economy, standing for the public good and the private good each individual can get. A particular element from $X$ will be denoted by $x = (y, x_1, x_2)$ . We will identify $X_1 \equiv X_2 \equiv E_+$ for simplicity and consider a fixed finite quantity $\overline{x} > 0$ of the private good to be distributed among the individuals. Let $T = \{(x_1, x_2) \in X_1 \times X_2 \ s.t. \ x_1 + x_2 \leq \overline{x}\}$ be the feasibility constraint on the private good, so assuming that both goods are technologically independent, the set of feasible alternatives will be: $FA \equiv X \cap (Y \times T)$ . Both agents are endowed with preferences representable by a continuous utility function: $u_i : Y \times X_i \rightarrow E$. for $i = 1, 2$, defined on his affective space - in Hurwicz &Walker's [11] terminology -. We will suppose the functions to be strictly increasing in the quantity of the private good and not quasi-linear: $\forall i = 1, 2, \ \forall y, \widehat{y} \in Y, \ u_i(y, x_i) - u_i(\widehat{y}, x_i) = u_i(y, \widetilde{x}_i) - u_i(\widehat{y}, \widetilde{x}_i) \Rightarrow x_i = \widetilde{x}_i$. The set of these admissible preferences will be $\Theta_i, \ i = 1, 2$. The reason why we completely exclude quasi-linear preferences on the domain is that Hurwicz &Walker [11] proved a similar theorem only for this kind of preferences. If we admit them into our new domain, the impossibility result becomes trivial and has no interest.

**Proposition 63** *Consider the above economy. There does not exist any in-*

*centive compatible compensation-free mechanism implementing the Pareto-optimal correspondence.*

**Proof.** First of all, notice that we do not need to focus on the whole Pareto-optimal correspondence since incentive compatibility requires that every selection correspondence must satisfy it. Let us take the following selection correspondence of the Pareto-optimal rule:
$\widehat{K}^*(\theta_1,\theta_2) = \underset{x \in FA}{\arg\max}\ v_1(y,x_1,\theta_1) + v_2(y,x_2,\theta_2)$ . Now we will consider a quite narrower subdomain of characteristics belonging to $\Theta_i\ \forall i = 1,2$, which will be denoted by $\widehat{\Theta}_i\ \forall i = 1,2$. $\widehat{\Theta}_i = \{a_i, b_i \in E,\ a_i > 0\}$ and $v_i(y,x_i,\theta_i) = a_i x_i^2 y - \frac{b_i}{2} y^2$. Notice that $\widehat{\Theta}_i \subset \Theta_i\ \forall i = 1,2$, since all of them are continuous, strictly increasing on the private good and no quasi-linear. Hence, incentive compatibility will hold within this subdomain too. But we can prove that $\widehat{K}^*$ is individually decisive for both agents. First, notice that since the utility functions are strictly increasing in the private good, every allocation in the whole Pareto-optimal SCC -and, of course, any selection from this- distributes the total amount available of the good among the agents, so if agent 1 can secure any $y \in Y$ and any amount $0 < x_1 < \overline{x}$ for himself by declaring an appropriate type, he can actually select some Pareto-optimal alternative, since $x_2 = \overline{x} - x_1$ in any efficient allocation. Thus, we will prove the following, i.e., for individual 1, $\forall y \in Y,\ x_1 \in X_1\ \&\ x_1 \leq \overline{x}$ , $\forall \theta_2 \in \widehat{\Theta}_2,\ \exists \widehat{\theta}_1 \in \widehat{\Theta}_1 \ni (y, x_1, \overline{x} - x_1) \in \widehat{K}^*(\widehat{\theta}_1, \theta_2) \in FA$. We can easily find the $\widehat{K}^*$ correspondence[17] $\forall(\theta_1,\theta_2) \in \widehat{\Theta}_1 \times \widehat{\Theta}_2$ : $\widehat{x}_1^* = \left(\frac{a_2}{a_1+a_2}\right)\overline{x}$, $\widehat{x}_2^* = \left(\frac{a_1}{a_1+a_2}\right)\overline{x}$, $\widehat{y}^* = \begin{cases} \max\left\{\frac{a_1 a_2 \overline{x}^2}{(a_1+a_2)(b_1+b_2)},\ \underline{y}\right\} & if\ (b_1+b_2) > 0 \\ \overline{y} & otherwise \end{cases}$

Now, take individual 1: $\forall a_2 > 0,\ \forall b_2 \in E,\ \forall x_1\ s.t.\ \overline{x} > x_1 > 0,\ \forall y \in Y,\ \exists \widehat{a}_1 = a_2\left(\frac{\overline{x} - x_1}{x_1}\right) > 0$,

$\exists \widehat{b}_1 = \frac{\widehat{a}_1 a_2 \overline{x} - (\widehat{a}_1 + a_2) b_2 y}{(\widehat{a}_1 + a_2)} \in E,\ s.t.\ \widehat{x}_1^*(\widehat{a}_1, a_2) = x_1\ \&\ \widehat{y}^*(\widehat{a}_1, a_2, \widehat{b}_1, b_2) = y$. The fact that agent Each agent cannot achieve the extremes - everything

[17] Note that $PO(\theta)$ in this case requires some positive amount of the private good to be given to both agents, so that this SCC can be shown to be individually decisive for the set $K$ restricted to $x_i > 0\ \forall i \in \{1,2\}$. To apply our result to the whole closed set $K$ we need some continuity assumptions, which is the problem we will deal with in the Appendix.

or nothing - of the total endowment of the private good prevent it to be individually decisive in the whole $K$, but it is clear that he can get quantities of the private good as close as desired to both extremes and since we are going to prove an impossibility.

Henceforth, applying *Theorem 58* and taking into account that we only admit compensation-free mechanisms, we have that agent $1's$ valuation function should be of the form: $v_1(g(\widehat{\theta}_1, \widehat{\theta}_2), \theta_1) = W(g(\widehat{\theta}_1, \widehat{\theta}_2), \theta_1, \widehat{\theta}_2),\ \forall\theta_1, \widehat{\theta}_1 \in \widehat{\Theta}_1,\ \forall\widehat{\theta}_2 \in \widehat{\Theta}_2$, for any selection $g$ from $\widehat{K}^*$. Since $v_1(y, x_1, \theta_1) + v_2(y, x_2, \theta_2)$ is a twice differentiable function -and concave for large sets of parameters-, the following equation has to hold for any $W \in \overline{W}(\widehat{K}^*)$ within those range of parameters:

$$\frac{\partial W(x_1, y, a_1, b_1, \widehat{b}_1, \widehat{b}_2)}{\partial y}(x_1^*, y^*) = 0 = \frac{\partial v_1(x_1, y, a_1, b_1)}{\partial y}(x_1^*, y^*).$$

And some simple calculations show that $\frac{\partial W(x_1, y, a_1, b_1, \widehat{b}_1, \widehat{b}_2)}{\partial y}(x_1^*, y^*) = a_1 x_1^{*2} - b_1 y^* - \widehat{b}_2 y^*$

and $\frac{\partial v_1(x_1, y, a_1, b_1)}{\partial y}(x_1^*, y^*) = a_1 x_1^{*2} - b_1 y^* \neq a_1 x_1^{*2} - b_1 y^* - \widehat{b}_2 y^*$ for all $\widehat{b}_2 \neq 0$, so the impossibility is proved. ■

The former proposition provides strong evidence about the non-existence of truthful revelation mechanisms for mixed economies that both generically provide the efficient quantities of the public and private goods and balance the budget when we limit ourselves to domains of preferences or characteristics with some income effect, so the income effect cannot be generically used to enforce strong implementation. Nevertheless, slightly different domains might lead to different implementation results, so the general question of existence is still open. Notice, for example, that the former proof is not valid with the additional restriction of concavity imposed on the preferences for the public good.

Moreover, when more than 2 agents are present, we may find that in some contexts, the SCC is not individually decisive, but it is for some fixed subset of the feasible alternatives -for the individual consumption spaces, but not for the others' consumptions-. In those cases, although we cannot directly apply the results in the former section, we can refine them to get useful tools to deal with those problems. For example, the following straightforward lemma might be applied:

**Lemma 64** *Let $K_i$, $\forall i \in N$, be a family of compact and convex sets. Let $v_i : K_i \times \Theta_i \to E$ be the continuous valuations for each individual and characteristic and suppose $K_i^*(\theta_1, .., \theta_n)$ be the restriction of $K^* : \prod_{i=1}^n \Theta_i \to \cup_{i=1}^n K_i$ on $K_i$. If $K_i^*$ is individually decisive on $K_i$ for some $i$, the only compensation-free mechanism that implements $K^*$ should be such that: $\forall \theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$, $v_i(g_i(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i) = W(g_i(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i})$, and for any selection $g$ from $K^*$. (where $g_i$ stands for the restriction of any $g$ on $K_i$).*

The proof of this lemma is omitted since it follows strictly the same reasoning of that in *Theorem 58* with the difference of considering $K_i$ as the whole set $K$ in the proof.

We can explore now the efficient implementation problem with compensation-free mechanisms in the extreme cases of pure public goods and private goods. Consider the public good case: depending on the admissible domain of characteristic we may choose, we can easily check if the Pareto-optimal SCC is individually decisive. For example, if we focus on the unrestricted domain of characteristics -see *Example 53* in the first section-, we can apply *Corollary 62* to get a strong version of Gibbard-Satterthwaite impossibility result. For the case of private goods, the strategy is similar: Let us consider, for example, the 2-agents, 2-goods general environment in *Example 54*, which is the classical Edgeworth Box economy, where the admissible characteristics are all continuous, strictly monotonic and convex utility functions over the 2-goods commodity space $E_+^2$. If we want to implement the Pareto-optimal SCC by means of compensation-free mechanisms, and it has been shown to be an individually decisive SCC, we can apply *Theorem 58* and get a similar impossibility to that in *Proposition 63* for the mixed economy close to that of Hurwicz [10]. Nevertheless, there is a difference that we should point out: Our work always assumes **full implementation**, since every selection of the SCC has to yield every individual's highest payoff, so when the set of possible social choices is large -as in the case of private goods along the contract curve-, it is easier to get an impossibility result and there might be selections of the SCC that can be implemented by revelation in a partial implementation framework-. As an example, consider the Pareto-optimal SCC when individuals are endowed with single-peaked preferences in *Example 55*'s economy in *Section 2*: The whole Pareto-optimal SCC with that domain is individually decisive, so it is easy to prove that there do not exist compensation-free mechanisms implementing it, but there are selections from this characterized by Moulin [12] -the SCF in the example is one of them- such that are not individually decisive and can be implemented by revelation.

Finally, we should note that the compensation mechanisms reproducing social objectives can be useful even in the absence of individual decisiveness. For some quasi-linear preferences in public goods environments, it is possible to achieve complete efficiency in the mixed economy using some particular Groves' mechanism if we restrict the domain even more. Groves and Loeb [9] found that the quadratic family of valuation functions on some public good in *Example 57* in *Section 2*, joint with quasi-linear preferences on the private good can be balanced with an appropriate Groves' mechanism, so that the efficient choice of the public good can be implemented by revelation for that domain. Notice, however, that this result does not contradict *Corollary 61*, since balanced implementation requires some appropriate compensation functions (transfers) and only this selection of the Pareto-optimal rule is implemented. Since a part of these transfers only depends on the others' reported types, no individual can get any transfer irrespective of the others' strategies, so the implementable Pareto-optimal selection is not individually decisive in the part of the private good so *Corollary 61* does not apply.

### 4.4.2 Full compensation scheme

**Corollary 65** *When the payoff function structure is that of* full compensation*, i.e.,* $\forall x, y_i \in E,\ P_i(x, y_i) = y_i,$ *every SCC can be trivially implemented by revelation by means of some compensation mechanism.*

**Proof.** Consider an incentive compatible compensation mechanism with full compensation payoff functions for every agent; it is easy to check that these payoff functions are monotonic, so applying *Theorem 58*, every individually decisive SCC implementable by revelation is such that: $\forall i \in N$, $\forall \theta_i, \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j,$

$W(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i})) = P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})$. Notice that the implementable SWF admitted in the full compensation case cannot depend on the real individual characteristics, so every SWF allowed is such that: $W(k, \theta) = f(k)\ \forall \theta \in \prod_{i=1}^{n} \Theta_i$, where $f : K \to E$ is any function. But, what kind of SCCs are represented by such SWFs? The only class is the following: $\exists \overline{K} \subseteq K$ such that $\forall \theta \in \prod_{i=1}^{n} \Theta_i,\ K^*(\theta) = \overline{K}$. Even though many members in that class are trivial undesirable SCCs like all the cases where $\overline{K}$ is a singleton, the case $\overline{K} \equiv K$ includes every possible social choice function as a selection from $K^*$. Moreover, we can say more about the form of the compensation functions: since $q_i(\widehat{\theta}_i, \widehat{\theta}_{-i}) \in E\ \forall \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$, incentive compatibility implies the following: $\forall \theta_i, \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j,\ q_i(\theta_i, \widehat{\theta}_{-i}) \geq q_i(\widehat{\theta}_i, \widehat{\theta}_{-i}) \Rightarrow q_i(\theta_i, \widehat{\theta}_{-i}) \geq range \left\{ q_i(\widehat{\theta}_i, \widehat{\theta}_{-i}) \right\} \Rightarrow$

$$\Rightarrow q_i(\theta_i, \widehat{\theta}_{-i}) \geq \max_{(\widehat{\theta}_i, \widehat{\theta}_{-i}) \in \prod_{i=1}^{n} \Theta_i} q_i(\widehat{\theta}_i, \widehat{\theta}_{-i}) \quad \forall \theta_i \in \Theta_i, \Rightarrow \forall \theta_i, \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j,$$

$q_i(\theta_i, \widehat{\theta}_{-i}) = q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})$, or, in other words, $\forall i \in N,\ \forall \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j,\ q_i(\widehat{\theta}_i, \widehat{\theta}_{-i}) = q_i(\widehat{\theta}_{-i})$. The compensation mechanisms associated to those compensation functions with the full compensation payoff function structure are independent of the own reported type, so they are always incentive compatible and (trivially) implement every SCC -not only the individually decisive SCCs-. Notice that for this trivial compensation functions, *Theorem 58* assigns the constant SWF, which trivially represents any SCC. ■

Notice that both the compensation-free and full compensation schemes are extreme or polar compensation mechanisms with opposite implementation properties: The impossibility of compensations makes the agents' payoff fully depending on their characteristics, so they are strongly interested in exploiting their private information, while if the agents' payoff can be designed completely independent of their types, any compensation scheme for an agent such that makes no use of his reported private information works. The reason is that in the full compensation case, the planner (or principal) owns the total power to modify the agent's payoff against changes in the characteristics, while in the no compensation scheme he can only use his discretion about the selected alternative, and finally there will only be strong implementation possibilities if the planner himself behaves as a dictatorial agent. Nevertheless, this trivial case has a clearly undesirable property: the agents have no incentives to lie, but they have not an incentive to tell the truth either. For a discussion of a similar setup, see Groves [7].

The remainder will study some specific intermediate cases between the full compensation and the no compensation possibilities.

### 4.4.3 Compensations with transfers

**Proposition 66** *Let $K$ be some compact set in a topological space and $\Theta_i$ be the set of all upper semi-continuous functions $\forall i \in N$, the only incentive compatible compensation mechanisms that implement the SWF $W = \sum_{i=1}^{n} v_i(k, \theta_i)$ are the following:*

$$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) =$$

$$= W \left[ \begin{array}{l} \\ \arg\max \quad v_i(k, \theta_i) + \sum_{j \neq i} v_j(k, \widehat{\theta}_j) \\ k \in K \end{array} \right].$$

$\forall i \in N,\ \forall \theta_i, \widehat{\theta}_i \in \Theta_i,\ \forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$, *where $W$ is any SWF representing the SCC in brackets.*

**Proof.** It suffices to prove that under the above condition, the SWF $W$ is individually decisive.

Take any $i \in N$ and any $\theta'_{-i} \in \prod_{j \neq i} \Theta_j$; as every $\theta'_j \in \Theta_j \quad \forall j \in N$, is bounded above by assumption, then, for any $\overline{k} \in K$, take the following type for individual $i$:

$v_i(k, \widehat{\theta}_i) = \begin{cases} -\sum_{j \neq i} v_j(k, \theta'_j) + 1 & if \ \ k = \overline{k} \\ -\sum_{j \neq i} v_j(k, \theta'_j) & if \ \ k \neq \overline{k} \end{cases}$ which is clearly upper semi-continuous and it will be true that:

$\forall \theta'_{-i} \in \prod_{j \neq i} \Theta_j, \quad W(\overline{k}, \widehat{\theta}_i, \theta'_{-i}) = v_i(\overline{k}, \widehat{\theta}_i) + \sum_{j \neq i} v_j(\overline{k}, \theta'_j) > v_i(k, \widehat{\theta}_i) + \sum_{j \neq i} v_j(k, \theta'_j) = W(k, \widehat{\theta}_i, \theta'_{-i})$

$\forall k \neq \overline{k} \in K$. This clearly implies that $\overline{k} \in K^*(\widehat{\theta}_i, \theta'_{-i})$, so for any $\forall \overline{k} \in K \ \ \forall \theta'_{-i} \in \Theta_{-i} \ \ \exists \widehat{\theta}_i \in \Theta_i$ s.t. $\overline{k} \in K^*(\widehat{\theta}_i, \theta'_{-i})$, so the SWF is individually decisive, because we can do the same for any $i \in N$.

Then, we are under the conditions of *Theorem 58*, so the only compensation mechanisms implementing by revelation $W$ is of the form described above.

**Proposition 67** *Let $K$ be some compact set in a topological space and the set $\Theta_i$ be the set of all continuous functions $\forall i \in N$, the only incentive compatible compensation mechanisms that implement the SWF: $W = \sum_{i=1}^{n} v_i(k, \theta_i)$, are the same of* Proposition 66*.*

**Proof.** We has to show that the SWF $W$ is individually decisive even when we restrict the domain of types to be continuous. Consider, then, for any $i \in N$, and for any $\overline{k} \in K$, given any $\theta'_{-i} \in \prod_{j \neq i} \Theta_j$, the following characteristic: $v_i(k, \widehat{\theta}_i) = -z \left\| k - \overline{k} \right\| - \sum_{j \neq i} v_j(k, \theta'_j) - \widehat{\varepsilon}(k)$. $z \in E_+$, & $\widehat{\varepsilon}(k)$ being any continuous function such that: $\widehat{\varepsilon}(k) \geq 0 \;\; \forall k \in K$ & $\widehat{\varepsilon}(\overline{k}) = 0$. We can prove that the SWF. with the profile $(\widehat{\theta}_i, \theta'_{-i})$ get a maximum on $k = \overline{k}$ :

$W(\overline{k}, \widehat{\theta}_i, \theta'_{-i}) = -\sum_{j \neq i} v_j(\overline{k}, \theta'_j) + \sum_{j \neq i} v_j(\overline{k}, \theta_j) - \widehat{\varepsilon}(\overline{k}) = -\widehat{\varepsilon}(\overline{k}) > -\widehat{\varepsilon}(k) - z \left\| k - \overline{k} \right\| =$

$-\sum_{j \neq i} v_j(k, \theta'_j) + \sum_{j \neq i} v_j(k, \theta'_j) - \widehat{\varepsilon}(k) - z \left\| k - \overline{k} \right\| = W(k, \widehat{\theta}_i, \theta'_{-i}) \;\; \forall k \in K$.

Notice that $v_i(k, \widehat{\theta}_i)$ is continuous since every $v_j(k, \theta'_j) \; \forall j \neq i$ are continuous, so we conclude as before, that $W$ is individually decisive. Applying *Theorem 58*, we can get the Generalized Groves' mechanisms again, being this a second generalization of Green & Laffont's results. ■

**Proposition 68** *Let $K \subset E^m$ be an open set endowed with the euclidean metric for $m \geq 2$ and the set $\Theta_i$ be the set containing all concave or strictly concave and differentiable functions $\forall i \in N$; the only incentive compatible compensation mechanisms that implement the SWF $W = \sum_{i=1}^{n} v_i(k, \theta_i)$ are the same that those in* Proposition 66*.*

**Proof.** We will show that $W$ has to be individually decisive even when the domain of characteristics is restricted to be every concave -or strictly concave- and differentiable function: First, consider any $i \in N$, and for any $\overline{k} \in K$, given any $\theta_{-i} \in \prod_{j \neq i} \Theta_j$, since every $v_j(k, \theta_j) \; \forall j \in N$ is a differentiable function, the expression: $\sum_{j \neq i}(v_j(k, \theta_j) - v_j(\overline{k}, \theta_j))$ is differentiable, so we know that there exists a vector $l(\overline{k}, \theta_{-i}) \in E^m$, $-\infty < l(\overline{k}, \theta_{-i}) < \infty$, such that:

$$lim_{\left\| k - \overline{k} \right\| \to 0} \frac{\left| \sum_{j \neq i}(v_j(k, \theta_j) - v_j(\overline{k}, \theta_j)) - l(\overline{k}, \theta_{-i})'(k - \overline{k}) \right|}{\left\| k - \overline{k} \right\|} = 0.$$

Now, let us construct a real number $\widehat{h}(\overline{k}, \theta_{-i}) \in E$ defined as follows:

$$\widehat{h}(\overline{k}, \theta_{-i}) = \\ = \max_{k \in K} \begin{cases} \dfrac{\left| \sum_{j \neq i}(v_j(k, \theta_j) - v_j(\overline{k}, \theta_j)) - l(\overline{k}, \theta_{-i})'(k - \overline{k}) \right|}{\left\| k - \overline{k} \right\|} & if \; k \neq \overline{k}. \\ 0 & if \; k = \overline{k} \end{cases}$$

Again, this number will exist because all the $v_j(k, \theta_j)$ are always bounded above and the denominator is positive, so $0 < \widehat{h}(\overline{k}, \theta_{-i}) < \infty$. Finally, we can write the following by construction:

$$\widehat{h}(\overline{k},\theta_{-i}) \geq \frac{\left|\sum_{j\neq i}(v_j(k,\theta_j)-v_j(\overline{k},\theta_j))-l(\overline{k},\theta_{-i})'(k-\overline{k})\right|}{\left\|k-\overline{k}\right\|} \quad \forall k\in K.$$

Rearranging the above inequality, we will have:

$\widehat{h}(\overline{k},\theta_{-i})\left\|k-\overline{k}\right\| \geq \left|\sum_{j\neq i}(v_j(k,\theta_j)-v_j(\overline{k},\theta_j))-l(\overline{k},\theta_{-i})'(k-\overline{k})\right| \geq$

$\geq \sum_{j\neq i}(v_j(k,\theta_j)-v_j(\overline{k},\theta_j))-l(\overline{k},\theta_{-i})'(k-\overline{k}). \Rightarrow$

$\widehat{h}(\overline{k},\theta_{-i})\left\|k-\overline{k}\right\| - \sum_{j\neq i}v_j(k,\theta_j) \geq -\sum_{j\neq i}v_j(\overline{k},\theta_j) - l(\overline{k},\theta_{-i})'(k-\overline{k})$

$\forall k\in K,$ and finally, multiplying the inequality by $-1$, and rearranging terms, we have: $\sum_{j\neq i}v_j(\overline{k},\theta_j) \geq -\widehat{h}(\overline{k},\theta_{-i})\left\|k-\overline{k}\right\| - l(\overline{k},\theta_{-i})'(k-\overline{k}) + \sum_{j\neq i}v_j(k,\theta_j)$

$\forall k\in K$. (1). So, we proved that $\forall i\in N,\ \forall\overline{k}\in K,\ \forall\ \theta_{-i}\in\prod_{j\neq i}\Theta_j,$

$\exists\widehat{h}(\overline{k},\theta_{-i})<\infty$ such that the last expression holds for all $k\in K$. Now, define the following characteristic for individual $i$:

$v_i(k,\widetilde{\theta}_i) = -\widetilde{h}(\overline{k},\theta_{-i})\left\|k-\overline{k}\right\| - l(\overline{k},\theta_{-i})'(k-\overline{k})$. Notice that $v_i(\overline{k},\widetilde{\theta}_i) = 0$ and it is easy to see that $\widetilde{\theta}_i\in\Theta_i$, since $\widetilde{h}(\overline{k},\theta_{-i})$ and $l(\overline{k},\theta_{-i})$ exist and it is a differentiable function - it is the sum of two differentiable functions - and is concave since the euclidean norm is strictly convex and the second term is convex. ($\|\lambda x+(1-\lambda)y\| \leq \|\lambda x\|+\|(1-\lambda)y\| = |\lambda|\,\|x\|+|(1-\lambda)|\,\|y\|\ \forall\lambda\in[0,1],\quad \forall x,y\in K.$ ). Now, the only thing to do is interpreting expression (1) as follows:

$\forall\overline{k}\in K,\ \forall\theta_{-i}\in\prod_{j\neq i}\Theta_j,\quad \exists\widetilde{h}(\overline{k},\theta_{-i})<\infty,\ \exists l(\overline{k},\theta_{-i})\in E^n,\ \exists\widetilde{\theta}_i\in\Theta_i$ such that: $v_i(\overline{k},\widetilde{\theta}_i)+\sum_{j\neq i}v_j(\overline{k},\theta_j) = 0+\sum_{j\neq i}v_j(\overline{k},\theta_j) \geq$

$\geq -\widetilde{h}(\overline{k},\theta_{-i})\left\|k-\overline{k}\right\| - l(\overline{k},\theta_{-i})'(k-\overline{k}) + \sum_{j\neq i}v_j(k,\theta_j) =$

$= v_i(k,\widetilde{\theta}_i)+\sum_{j\neq i}v_j(k,\theta_j)$

$\forall k\in K \Rightarrow v_i(\overline{k},\widetilde{\theta}_i)+\sum_{j\neq i}v_j(\overline{k},\theta_j) \geq v_i(k,\widetilde{\theta}_i)+\sum_{j\neq i}v_j(k,\theta_j)\ \ \forall k\in K \Rightarrow$

$\forall i\in N,\ \forall\overline{k}\in K,\ \forall\theta_{-i}\in\prod_{j\neq i}\Theta_j,\ \exists\widetilde{\theta}_i\in\Theta_i$ such that $W(\overline{k},\widetilde{\theta}_i,\theta_{-i}) \geq W(k,\widetilde{\theta}_i,\theta_{-i}) \Rightarrow \overline{k}\in K^*(\widetilde{\theta}_i,\theta_{-i})$.

So there exists some characteristic in the admissible domain such that any individual can get any alternative for any others' characteristics, which is the definition of an individually decisive SCC. Applying *Theorem 58*, we find the same class of mechanisms as above. ■

**Corollary 69** *(Green & Laffont* [5]*). Let $K$ be a compact set in a topological space and $\Theta_i$ $\forall i\in N$ contain any upper semi-continuous or continuous or concave and differentiable valuation functions. The only compensation by* *transfers mechanisms that can implement by revelation the utilitarian SWF are the Groves' mechanisms.*

In any of the possible domains considered, the only compensation mechanisms that can implement by revelation the utilitarian SWF are of the same form using *Propositions 66* to 68, i.e.,

$$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = W\left[\begin{array}{l} \arg\max \quad v_i(k, \theta_i) + \sum_{j \neq i} v_j(k, \widehat{\theta}_j) \\ k \in K \end{array}\right].$$

$\forall i \in N$, $\forall \theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$, where $W$ is any SWF representing the SCC in brackets.

Now, we are imposing an additional restriction to the compensation mechanisms allowed: we are only interested in compensations by transfers, i.e., the payoff functions structure have the form: $\forall x, y_i \in E,\ P_i(x, y_i) = x + y_i$, so, applying *Propositions 66* to 68:

$$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i) + \widehat{q}_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \widehat{\theta}_{-i}) =$$

$$= W\left[\begin{array}{l} \arg\max \quad v_i(k, \theta_i) + \sum_{j \neq i} v_j(k, \widehat{\theta}_j) \\ k \in K \end{array}\right] =$$

$$= v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i) + f(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \widehat{\theta}_{-i})),$$

$\forall i \in N$, $\forall \theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$, where $f$ is some adequate function, so it suffices to prove that $\sum_{j \neq i} v_j(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \widehat{\theta}_j) + h_i(\theta_{-i}) =$

$= f(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \widehat{\theta}_{-i}))$ $\forall i \in N$, $\forall \widehat{\theta} \in \prod_{i=1}^{n} \Theta_i$, where $h_i(\theta_{-i})$ is any real valued function, which is the only functional form allowed -see Green & Laffont [5]-.

Notice that the Groves' mechanisms are a particular case of those compensation mechanisms in the case of interpreting the model as choosing some vector of public goods. Notice that in *Propositions 66* to 68, we did not assume the quasi-linearity of the final payoff function, which will be interpreted in our context as allowing only for compensations by transfers, which is a particular member of the family of payoff function structures, so this propositions are stronger than Green & Laffont's Theorem in the sense that concluding that the only form of the utility function on both private and public goods that allows for the implementability of the utilitarian SWF is exactly the domain imposed by the former authors, i.e., the quasi-linear preferences without income effect. Notice, also, that only under this restriction on the domain of extended preferences the SWF as representing the Pareto optimal SCC makes complete sense.

An important feature of the model that should be pointed out is that Green and Laffont's results, as well as ours, are extremely dependent on the **non-existence of a common fixed bound** on the types allowed in the domain. Notice that if the planner possess the additional information that the characteristics cannot be *too high*, the utilitarian SCC will not be individually decisive. Suppose, for example, that the domain of types is restricted to be all the bounded and continuous or upper semi-continuous functions such that $\forall i \in N,\ \forall \theta_i \in \Theta_i,\ \exists \overline{c}, \underline{c} \in E\ s.t.\ \forall k \in K,\ \underline{c} \leq v_i(k, \theta_i) \leq \overline{c}$. With this new restriction, for some $\theta_{-i} \in \prod_{j \neq i} \Theta_j$, it is impossible to find feasible types $\theta_i \in \Theta_i,\ s.t.\ \forall k \in K,\ k \in \underset{k \in K}{\arg\max} \ \sum_{j=1}^{n} v_j(k, \theta_j)$. Hence, with this new domain, the utilitarian SWF is not individually decisive and *Theorem 58* cannot be applied. Take, for example, $n = 3,\ \underline{c} = 0,\ \overline{c} = 1$, and $\forall j \neq i,\ \ v_j(k, \theta_j) = \begin{cases} 1 & for\ k \geq \overline{k}. \\ 0 & otherwise. \end{cases}$ Notice that for any $\theta_i \in \Theta_i$, the following holds: $W(\widehat{k}, \theta_i, \theta_{-i}) \geq 2 > 1 \geq W(\widetilde{k}, \theta_i, \theta_{-i})\ \forall \widehat{k} \geq \overline{k},\ \forall \widetilde{k} < \overline{k}$.

### 4.4.4 Compensations by means of prices

Consider the following particular problem: $N = \{1, 2\}$,

$K = \left\{(k_1, k_2) \in E_+^2 \ \ s.t.\ k_1 + k_2 = \overline{k}\right\}$, $\forall \theta_1 \in \Theta_1,\ \ v_1(k, \theta_1)$ is such that: $\forall (k_1, k_2), (k_1, \widehat{k}_2) \in K,\ \ v_1(k_1, k_2, \theta_1) = v_1(k_1, \widehat{k}_2, \theta_1)$ & $\forall \theta_i \in \Theta_i,\ \exists (k_1, k_2) \in K$ such that: $v_i(k_1, k_2, \theta_i) > 0,\ \forall i = 1, 2$.

$v_1(k_1, k_2, \theta_1)$ is continuous and strictly increasing in the first argument. Agent 2's characteristics will be of the same kind but permuting the arguments of the set $K$. We will write them $v_1(k_1, \theta_1)$ and $v_2(k_2, \theta_2)$.

**Proposition 70** *The only incentive compatible compensation mechanisms that implement by revelation the Nash SWF, i.e., $W = v_1(k, \theta_1) v_2(k, \theta_2)$, are such that:* $\forall i \in \{1, 2\},\ \forall \theta_i \in \Theta_i,\ \forall \theta_j \in \Theta_j,\ (j \neq i)$,

$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_j), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_j)) =$

$= W(g(\widehat{\theta}_i, \widehat{\theta}_j)), \theta_i, \widehat{\theta}_j) = W\left[\underset{k \in K}{\arg\max} \ \ v_1(k, \theta_1) v_2(k, \theta_2)\right]$, *where $W$ is any SWF such that $W \in \overline{W}(K^*)$.*

**Proof.** We are going to prove that in the above economy, the Nash bargaining SWF is, in fact, individually decisive. We show this for individual 1 and the proof for the other agent is, of course, symmetric.

Suppose any admissible type for individual 2: $\theta_2 \in \Theta_2$, take any alternative from the range of the Nash SCC, i.e., the open set:

$$\widehat{K} = \left\{ (k_1, k_2) \in E_+^2 \;\; s.t.\; k_1 k_2 \neq 0 \;\;\&\;\; k_1 + k_2 = \overline{k} \right\}.$$

Notice that we cannot pick up alternatives with some zero component. Let us call any of this $\widehat{k} = (\widehat{k}_1, \widehat{k}_2) = (\widehat{k}_1, \overline{k} - \widehat{k}_1) \in \widehat{K}$. Now, consider the following characteristic for individual 1: for any $\theta_2 \in \Theta_2$ declared by agent 2,

$v_1(k_1, \widehat{\theta}_1) = 2v_1(\widehat{k}_1, \widehat{\theta}_1) - \dfrac{v_1(\widehat{k}_1, \widehat{\theta}_1)}{v_2(\widehat{k}_2, \theta_2)} v_2(\overline{k} - k_1, \theta_2)$. $v_1(\widehat{k}_1, \widehat{\theta}_1)$ being any positive real number. Let us call $\widehat{\beta}(\widehat{\theta}_1, \theta_2)$ the absolute slope of the function $\widehat{\beta}(\widehat{\theta}_1, \theta_2) = \dfrac{v_1(\widehat{k}_1, \widehat{\theta}_1)}{v_2(\widehat{k}_2, \theta_2)}$. Notice that this is an admissible type for individual 1, because it is decreasing on $k_2$ (increasing on $k_1$) when $(k_1, k_2) \in \widehat{K}$, and continuous since $\theta_2$ is continuous. By using this strategy, the outcome of the SWF will be the set:

$$K^*(\widehat{\theta}_1, \theta_2) = \underset{\substack{k_1, k_2 \\ s.t}}{\arg\max} \;\; W = v_1(k_1, \widehat{\theta}_1) v_2(k_2, \theta_2) \quad k_1 + k_2 \leq \overline{k}$$

Suppose that $(k_1^*, k_2^*) \in K^*(\widehat{\theta}_1, \theta_2)$, it always holds that, if we define $\widehat{u}_1^*(k_1) = v_1(k_1, \widehat{\theta}_1)$, and $u_2^*(k_2) = v_2(k_2, \theta_2)$

$$(\widehat{u}_1^*(k_1^*), u_2^*(k_2^*)) = \underset{\substack{\widehat{u}_1, u_2 \\ s.t}}{\arg\max} \;\; W = \widehat{u}_1 u_2 \quad \widehat{u}_1 = 2v_1(\widehat{k}_1, \widehat{\theta}_1) - \widehat{\beta}(\widehat{\theta}_1, \theta_2) u_2$$

But the feasibility constraint is a linear function in the space $(\widehat{u}_1, u_2)$, so the solution will be unique and the necessary and sufficient conditions that hold in the optimum are the following:

i) $\widehat{\beta} = \dfrac{\widehat{u}_1(k_1^*)}{u_2(k_2^*)} = \dfrac{du_1}{du_2} = \widehat{\beta}(\widehat{\theta}_1, \theta_2) = \dfrac{v_1(\widehat{k}_1, \widehat{\theta}_1)}{v_2(\widehat{k}_2, \theta_2)}$ (By definition of the slope).

ii) $\widehat{u}_1(k_1^*) = 2v_1(\widehat{k}_1, \widehat{\theta}_1) - \widehat{\beta}(\widehat{\theta}_1, \theta_2) u_2(\overline{k} - k_1^*)$.

Notice that by construction, the first tangency condition can only be fulfilled in $k^* = \widehat{k}$, and the second holds for $k_1^* = \widehat{k}_1$ too. Therefore, agent 1 has always some admissible strategy so that he can get any alternative he wants with the exception of the extremes, but he can choose alternatives as closed as desired to them (an alternative out of the range will make the slope become infinity) and always any alternative in $range(K^*)$, so the SWF is individually decisive. Since the types are continuous functions, we can now apply *Theorem 58* and the conclusion is obvious. ■

This proposition posses its own interest since the above admissible domain of characteristics can be interpreted as production sets of two firms that have to share some fixed amount of a common input. Both firms have private information about his own technology and send their revealed technologic characteristics to the planner or authority, which makes the sharing final decision. It shows that we can implement by revelation the Nash bargaining solution even with lack of information if the authority can establish the prices at which both firms sell their respective output. Therefore, we should use some price payoff structure to get it.

**Corollary 71** *Suppose $\theta_1, \theta_2 \in \Theta_1, \Theta_2$. Every individually decisive SCC can be trivially implemented by means of prices, i.e., when individual's payoff functions are of the type:*
$P_i(v_i(g(\theta_i, \theta_{-i}), \theta_i), q_i(\theta_i, \theta_{-i})) = v_i(g(\theta_i, \theta_{-i}), \theta_i) q_i(\theta_i, \theta_{-i})$.

**Proof.** Obvious: just consider the following compensation functions: $q_i(\theta_i, \theta_{-i}) = 0 \; \forall \theta_i \in \Theta_i, \; \forall \theta_{-i} \in \prod_{j \neq i} \theta_j$. ■

### 4.4.5 Other types of compensations

Now, we may wonder what other SWFs the planner could be interested to implement. Thinking on some ethical and efficient rules, we can investigate if there exists a method to implement some kind of equal welfare among the agents.

**Proposition 72** *Let $K$ be a compact set in a topological space and $\Theta_i$ be the set of continuous or upper semi-continuous bounded functions. There does not exist any incentive compatible compensation mechanism implementing by revelation the Rawlsian SWF, i.e., $W = min\left\{v_1(k, \theta_1), v_2(k, \theta_2), ..., v_n(k, \theta_n)\right\}$.*

**Proof.** We prove that if the set of characteristics $\Theta_i$ $\forall i \in N$, are either upper semi-continuous or continuous, the Rawlsian or egalitarian SWF represent an individually decisive SCC. Suppose any individual $i$, and any $\theta_{-i} \in \prod_{j \neq i} \Theta_j$, set any $\overline{k} \in K$, and consider the real number $c_j \in E$ be the lower bound of each $j \neq i$, then, there exists the number $\widehat{c}_i = \min_{j \neq i} \{c_j\} < \infty$, and, for any $\overline{k} \in K$, take the following type for individual $i$:

$v_i(k, \widehat{\theta}_i) = \begin{cases} \widehat{c}_i & if \;\; k = \overline{k} \\ \widehat{c}_i - 1 & if \;\; k \neq \overline{k} \end{cases}$ Notice that this is a bounded below, upper semi-continuous function, and it holds by construction that:

$$\min \{v_1(k, \theta_1), ..., v_{i-1}(k, \theta_{i-1}), v_{i+1}(k, \theta_{i+1}), ..., v_n(k, \theta_n)\} \geq \min_{j \neq i} \{c_j\} = \widehat{c}_i,$$

because every characteristic is bounded below by $c_j$, so $\min_{j \neq i} \{v_j(k, \theta_j)\} \geq \widehat{c}_i \geq v_i(k, \widehat{\theta}_i)$ $\forall k \in K$. Consider now the problem:

$$\max_{k \in K} \quad \min \left\{ v_1(k, \theta_1), ..., v_i(k, \widehat{\theta}_i), ..., v_n(k, \theta_n) \right\} = \max_{k \in K} \quad v_i(k, \widehat{\theta}_i)$$

and observing the definition of $v_i(k, \widehat{\theta}_i)$, $v_i(\overline{k}, \theta_i) \geq v_i(k, \widehat{\theta}_i)$ $\forall k \in K$, so $\overline{k} \in K^*(k, \widehat{\theta}_i, \theta_{-i})$, and we can generate some $v_i(k, \widehat{\theta}_i)$ for every $\overline{k} \in K$, and every $\theta_{-i} \in \prod_{j \neq i} \Theta_j$, so the Rawlsian SCC is individually decisive. The case where $\Theta_i$ contains only continuous preferences does not differ very much from this one: just consider instead of $\widehat{\theta}_i \in \Theta_i$, the following characteristic: $\widetilde{\theta}_i \in \Theta_i$ such that:

$v_i(k, \widetilde{\theta}_i) = \begin{cases} \widehat{c}_i - \left\| k - \overline{k} \right\| & for \;\; k \;\; s.t. \;\; \left\| k - \overline{k} \right\| \leq 1 \\ \widehat{c}_i - 1 & otherwise. \end{cases}$ And it is easy to check that the function is continuous and individual $i$ can get any alternative he wants just by changing $\overline{k}$, $\forall k \in K$.

Now we can apply *Theorem 58* in both cases and obtain: $\forall i \in N$, $\forall \theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$,

$P_i(v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), q_i(\widehat{\theta}_i, \widehat{\theta}_{-i})) = W((\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i, \widehat{\theta}_{-i}) =$

$f(\min \left\{ v_i(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_i), v_{-i}(g(\widehat{\theta}_i, \widehat{\theta}_{-i}), \theta_{-i}) \right\})$, for any selection $g$ from $K^*$ and $f$ being some function. But notice that the payoff function structure associated with the mechanism has to be the following: $P_i(v_i(k, \theta_i), q_i(k, \widehat{\theta}_{-i})) = W \left[ \arg\min_k \left\{ v_i(k, \theta_i), q_i(k, \widehat{\theta}_{-i}) \right\} \right]$, for any $W \in W(K^*)$. But all these are non-monotonic payoff functions: Suppose $\theta \in \prod_{i=1}^n \Theta_i$ such that $\exists i \in N$ such that $v_i(k,\theta_i) < v_j(k,\theta_j)\ \forall k \in K$. Then, $K^*(\theta) = \underset{k \in K}{\arg\max}\ v_i(k,\theta_i)$ and the compensation cannot change. Suppose $\overline{q} \in E$, $\widehat{\theta}_i \in \Theta_i$ such that: $\overline{q} > v_i(k,\widehat{\theta}_i)$, $\forall k \in K$, then, if we consider $\widetilde{q} > \overline{q}$, $\widetilde{q} > \overline{q} > v_i(k,\widehat{\theta}_i)$, and the payoff remains the same: $P_i(v_i(k,\widehat{\theta}_i),\overline{\theta}) = P_i(v_i(k,\widehat{\theta}_i),\widetilde{\theta}) = v_i(k,\widehat{\theta}_i)$, a contradiction with the monotonicity assumption, so there cannot exist compensation mechanisms implementing the Rawlsian SWF under all the conditions above. ■

**Corollary 73** *The following compensation and non-monotonic mechanism allows for implementation of the Rawlsian SCC in the above economy:* $\forall \theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$,

$$P_i(v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\theta_i), q_i(\widehat{\theta}_i,\widehat{\theta}_{-i})) =$$
$$= \max \left\{ -v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\theta_i),\ -\underset{j \neq i}{\max} \left\{ v_j(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\widehat{\theta}_j) \right\} \right\}.$$

**Proof.** Notice that the expression above can be written as:

$$= -\min \left\{ v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\theta_i),\ \underset{j \neq i}{\max} \left\{ v_j(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\widehat{\theta}_j) \right\} \right\} =$$
$$= -\min \left\{ v_1(g(\widehat{\theta}),\theta_i), ..., v_i(g(\widehat{\theta}),\theta_i), ..., v_n(g(\widehat{\theta}),\widehat{\theta}_n) \right\}.$$

-of course the associated mechanism is not monotonic-. Abusing notation, we can write:

$$P_i(v_i(g(\theta_i,\widehat{\theta}_{-i}),\theta_i), q_i(\theta_i,\widehat{\theta}_{-i})) =$$
$$-\min \left\{ v_i(g(\theta_i,\widehat{\theta}_{-i}),\theta_i),\ v_{-i}(g(\theta_i,\widehat{\theta}_{-i}),\widehat{\theta}_{-i}) \right\} \geq$$
$$\geq -\min \left\{ v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\widehat{\theta}_i),\ v_{-i}(g(\widehat{\theta}_i,\widehat{\theta}_{-i}),\widehat{\theta}_{-i}) \right\} =$$
$$= P_i(v_i(g(\widehat{\theta}_i,\widehat{\theta}_{-i})), q_i(\widehat{\theta}_i,\widehat{\theta}_{-i}))$$

$\forall \theta_i, \widehat{\theta}_i \in \Theta_i$, $\forall \widehat{\theta}_{-i} \in \prod_{j \neq i} \Theta_j$, and this for any selection $g(\widehat{\theta})$ from $K^*(\widehat{\theta})$, so the non-monotonic compensation mechanism implements the Rawlsian SCC. ■

## 4.5 Concluding remarks

We have proved in this paper that when we are trying to implement by revelation any SCC *too sensitive to individual preferences*, we have to rely on individual compensation mechanisms that replicate the social objectives. This is the reason why the well-known Groves' mechanism works in quasi-linear domains of preferences, which can be viewed in terms of our model. Hence, it is not by chance that the transfer any individual receives takes the same functional form that the SWF we are trying to implement, but is a general feature that can be extended to different social welfare criteria and to different compensation schemes. Therefore, there exists a strong linkage between the compensation structure we allow in each case and the social welfare functions we can implement: Essentially, we need an additive transfer scheme -like taxes or subsidies- to implement the utilitarian SWF, some multiplicative prices scheme to get the bargaining Nash solution, and it is impossible within our assumptions to implement the egalitarian rule. When the planner cannot make compensations, implementation by revelation requires dictatorship or it is often impossible. In the opposite extreme, when the planner can expropriate the part of the agents' objective functions affected by the type and completely determine the final payoff, every social choice rule can be trivially implemented. It seems that diminishing the effect of the agents' types on their own payoff considerably enlarge the set of rules that can be implemented. We have abstracted thorough the paper the possible costs the planner may face in choosing one or another contract structure -the mechanism-, but if they exist, the planner might compare the implementation gains with the costs associated to each contract. Moreover, one or another social choice rule or compensation mechanism might be more appropriate in different contexts: public good provision, production implementation, etc., but we are always constrained by the menu provided by *Theorem 58*.

## 4.6 Appendix

Although the dependence of *Theorem 58* to the individually decisiveness assumption is clear, we can slightly relax the class of admitted SCCs if we restrict attention to continuous mechanisms (when both the compensation functions and the payoff functions are continuous when considering the sup norm). The following result provides the continuous mechanisms version of *Theorem 58.*, but, before we state it, we need one more definition and a lemma.

**Definition 74** *A SCC $K^*$ is called* ***individually quasi-decisive*** *if* $\forall i \in N$, $\forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$, $\forall \tau > 0$, $\forall k \in K$, $\exists \widehat{\theta}_i(\tau, k, \theta_{-i}) \in \Theta_i$ *such that* $\overline{k} \in K^*(\widehat{\theta}_i(\tau, k, \theta_{-i}), \theta_{-i}) \in K$, $\overline{k} \in B_\tau(k)$ & $\exists \lim_{\tau \to 0} \widehat{\theta}_i(\tau, k, \theta_{-i})$. *( where $B_\tau(k)$ stands for the open ball with center in $k$ and radius $\tau$ using the euclidean metric ).*

This property means that everybody can obtain an alternative as close as desired to any other by reporting an adequate type and is weaker than individual decisiveness -every individually decisive SCC is always quasi-decisive, but the converse is not true-.

**Lemma 75** *Suppose $\Theta_i$ contains only continuous functions for each $i$. Let $\{P, q\}$ be a continuous, incentive compatible for $i$ compensation mechanism implementing the SCC $K^*$, and take two -possibly the same- selections from $K^*$ : $g$ and $\widehat{g}$. Then, $\forall \theta_i, \theta_i' \in \Theta_i$, $\forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$, $\forall \phi > 0$, $\exists \rho > 0$ such that $g(\theta_i, \theta_{-i}) \in B_\rho(\widehat{g}(\theta_i', \theta_{-i})) \Rightarrow\ q_i(\theta_i, \theta_{-i}) \in B_\phi(q_i(\theta_i', \theta_{-i}))$.*

**Proof.** By contradiction, suppose that $\forall \rho > 0$, $\exists \theta_i, \theta_i' \in \Theta_i$, $\exists \theta_{-i} \in \prod_{j \neq i} \Theta_j$, $\exists \phi > 0$, such that $g(\theta_i, \theta_{-i}) \in B_\rho(\widehat{g}(\theta_i', \theta_{-i}))$ &
$q_i(\theta_i, \theta_{-i}) \notin B_\phi(q_i(\theta_i', \theta_{-i}))$. Suppose w.l.g. that
$q_i(\theta_i, \theta_{-i}) - q_i(\theta_i', \theta_{-i}) > \phi$. Then, take type $\theta_i'$, and consider the payoffs:
$P_i(v_i(g(\theta_i, \theta_{-i}), \theta_i'), q_i(\theta_i, \theta_{-i}))$ and $P_i(v_i(\widehat{g}(\theta_i', \theta_{-i}), \theta_i'), q_i(\theta_i', \theta_{-i}))$. But, by hypothesis, $\forall \rho > 0$, $\exists \theta_i, \theta_i' \in \Theta_i$, $\exists \theta_{-i} \in \prod_{j \neq i} \Theta_j$, such that $g(\theta_i, \theta_{-i}) \in B_\rho(\widehat{g}(\theta_i', \theta_{-i}))$, so we can choose $\theta_i, \theta_i'$ such that $g(\theta_i, \theta_{-i})$ will be as close to $\widehat{g}(\theta_i', \theta_{-i})$ as we want. Since $\Theta_i$ includes only continuous functions, we can make $v_i(g(\theta_i, \theta_{-i}), \theta_i')$ as close as desired to $v_i(\widehat{g}(\theta_i', \theta_{-i}), \theta_i')$ by choosing $\theta_i, \theta_i' \in \Theta_i$. But $P_i$ is continuous in both arguments and monotonic so, by sufficiently choosing $\rho > 0$ and $\theta_i, \theta_i' \in \Theta_i$, we should have, since $q_i(\theta_i, \theta_{-i}) - q_i(\theta_i', \theta_{-i}) > \phi$ :
$P_i(v_i(g(\theta_i, \theta_{-i}), \theta_i'), q_i(\theta_i, \theta_{-i})) > P_i(v_i(\widehat{g}(\theta_i', \theta_{-i}), \theta_i'), q_i(\theta_i', \theta_{-i})). \Rightarrow$
For $\theta_i'$, $\exists \theta_i \in \Theta_i$, $\exists \overline{g}(\theta_i, \theta_{-i}) \in K^*(\theta_i, \theta_{-i})$ $\forall \theta_i \in \Theta_i$, $\forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$, -a selection from $K^*$- defined as:

$$\overline{g}(\widehat{\theta}_i, \theta_{-i}) = \begin{cases} g(\widehat{\theta}_i, \theta_{-i}) & iff\ \widehat{\theta}_i \neq \theta_i' \\ \widehat{g}(\theta_i', \theta_{-i}) & iff\ \widehat{\theta}_i = \theta_i' \end{cases}$$

, such that by declaring $\theta_i \in \Theta_i$, agent $i$ will receive a greater payoff than reporting the truth $(\theta_i')$, i.e.,
$P_i(v_i(\overline{g}(\theta_i, \theta_{-i}), \theta_i'), q_i(\theta_i, \theta_{-i})) > P_i(v_i(\overline{g}(\theta_i', \theta_{-i}), \theta_i'), q_i(\theta_i', \theta_{-i}))$, so the mechanism cannot be incentive compatible for $i$, a contradiction. ■

**Theorem 76** *Suppose that $\forall i \in N,\ \Theta_i$ is restricted to contain continuous functions and let $K^*$ be an individually quasi-decisive SCC. Then if $\{P, q\}$ is an incentive compatible and continuous compensation mechanism that implement $K^*$, then it must be of the form of those in* Theorem 58*.*

**Proof.** $\Rightarrow$) First of all, notice that *Lemma 59* applies to every compensation mechanism implementing any SCC, so it holds now that $\forall \theta_i, \theta_i' \in \Theta_i,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$, such that $g(\theta_i, \theta_{-i}) = g(\theta_i', \theta_{-i})$, it holds that: $q_i(\theta_i, \theta_{-i}) = q_i(\theta_i', \theta_{-i})$.

Moreover, since $K^*$ is individually quasi-decisive, $\forall i \in N,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$, $\forall \tau > 0,\ \forall k \in K,\ \exists \widehat{\theta}_i : E_{++} \times K \times \prod_{j \neq i} \Theta_j \rightarrow\rightarrow \Theta_i$ such that $\overline{k} \in K^*(\widehat{\theta}_i(\tau, k, \theta_{-i}), \theta_{-i}) \in K$ & $\overline{k} \in B_\tau(k)$. Let us now define the following correspondence: $\widehat{q}_i : K \times \Theta_{-i} \rightarrow\rightarrow E$, such that: $\widehat{q}_i(k, \theta_{-i}) =$ $= q_i(\lim_{\tau \to 0} \widehat{\theta}_i(\tau, k, \theta_{-i}), \theta_{-i})$ for all $k \in K,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$.

This mapping is similar to the analogous one used in the proof of *Theorem 58*. Moreover, $\widehat{q}_i(k, \theta_{-i})$ always exist and is a **continuous function on** $K$. Notice that $\forall k \in K$ such that $\exists \theta_i' \in \Theta_i$ such that $k = g(\theta_i', \theta_{-i})$ for some selection $g$ from $K^*$, $\widehat{q}_i(k, \theta_{-i})$ is a singleton following the same reasoning in proving (ii) in *Theorem 58* -using *Lemma 59* for all $i$-, and the cases where $\forall \theta_i' \in \Theta_i,\ k \neq g(\theta_i', \theta_{-i})$, individual quasi-decisiveness assures that the limit exists. It remains to prove that the function is continuous for every $k \in K$, so we will prove the following: $\forall \theta_{-i} \in \prod_{j \neq i} \Theta_j,\ \forall k \in K,\ \forall \varepsilon > 0,\ \exists \delta(k, \varepsilon) > 0$ such that $\overline{k} \in B_\delta(k) \Rightarrow \widehat{q}_i(\overline{k}, \theta_{-i}) \in B_\varepsilon(\widehat{q}_i(k, \theta_{-i}))$. Suppose the contrary: $\exists \theta_{-i} \in \prod_{j \neq i} \Theta_j,\ \exists k \in K,\ \exists \varepsilon > 0$, such that $\forall \delta > 0,\ \exists \overline{k} \in B_\delta(k)$ & $\widehat{q}_i(\overline{k}, \theta_{-i}) \notin B_\varepsilon(\widehat{q}_i(k, \theta_{-i}))$. Then, choose any selection $\widehat{q}_i'(k, \theta_{-i}) \in \widehat{q}_i(k, \theta_{-i}),\ \forall k \in K,\ \forall \theta_{-i} \in \prod_{j \neq i} \Theta_j$ and set any $\theta_{-i} \in \prod_{j \neq i} \Theta_j$; Now, take any $k \in K$ such that $\exists \theta_i' \in \Theta_i$ such that $k = g(\theta_i', \theta_{-i})$ for some selection $g$ from $K^*$. Then, suppose that $\exists k \in K$ such that $\exists \varepsilon > 0$ such that $\forall \delta > 0,\ \exists \overline{k}(\delta) \in B_\delta(k)$ & $\widehat{q}_i'(\overline{k}, \theta_{-i}) \notin B_\varepsilon(\widehat{q}_i'(k, \theta_{-i}))$.

First, since $K^*$ is individually quasi-decisive, we know that there exists some selection $g \in K^*$ such that, for $\theta_{-i} \in \prod_{j \neq i} \Theta_j,\ k, \overline{k}(\delta) \in K$ and $\forall \delta > 0$, These two conditions hold:

(i). $\forall \tau' > 0,\ \exists \widetilde{\theta}_i \in \Theta_i$ such that $g(\widetilde{\theta}_i, \theta_{-i}) \in B_{\tau'}(k)$.

(ii). $\forall \tau > 0,\ \exists \overline{\theta}_i \in \Theta_i$ such that $g(\overline{\theta}_i, \theta_{-i}) \in B_\tau(\overline{k}(\delta))$.

Moreover, by definition of $\widehat{\theta}_i(\tau, k, \theta_{-i})$, we know that:

$\widetilde{\theta}_i \in B_{\tau'}(\lim_{\tau' \to 0} \widehat{\theta}_i(\tau', k, \theta_{-i}))$ and $\overline{\theta}_i \in B_\tau(\lim_{\tau \to 0} \widehat{\theta}_i(\tau, \overline{k}(\delta), \theta_{-i}))$ so it holds that $\forall \delta, \tau, \tau' > 0,\ \exists \widetilde{\theta}_i, \overline{\theta}_i \in \Theta_i$ such that $g(\widetilde{\theta}_i, \theta_{-i}) \in B_{\delta+\tau+\tau'}(g(\overline{\theta}_i, \theta_{-i}))$.

Furthermore, by continuity of the bounded compensation functions in $\lim_{\tau\to 0}\widehat{\theta}_i(\tau,\overline{k}(\delta),\theta_{-i})$ and $\lim_{\tau'\to 0}\widehat{\theta}_i(\tau',k,\theta_{-i})$, it is true that $\forall\widehat{\varepsilon}_1,\widehat{\varepsilon}_2>0$, $\exists\widehat{\delta}_1(\widehat{\varepsilon}_1),\widehat{\delta}_2(\widehat{\varepsilon}_2)>0$, such that $\theta_i\in B_{\widehat{\delta}_1(\widehat{\varepsilon}_1)}(\lim_{\tau\to 0}\widehat{\theta}_i(\tau,\overline{k}(\delta),\theta_{-i}))$ and $\theta_i'\in B_{\widehat{\delta}_2(\widehat{\varepsilon}_2)}(\lim_{\tau'\to 0}\widehat{\theta}_i(\tau',k,\theta_{-i})),\Rightarrow$
$\Rightarrow q_i(\theta_i,\theta_{-i})\in B_{\widehat{\varepsilon}_1}(q_i(\lim_{\tau\to 0}\widehat{\theta}_i(\tau,\overline{k}(\delta),\theta_{-i}),\theta_{-i}))$ and $q_i(\theta_i',\theta_{-i})\in B_{\widehat{\varepsilon}_2}(q_i(\lim_{\tau'\to 0}\widehat{\theta}_i(\tau',k,\theta_{-i}),\theta_{-i}))$. Now, given any $\delta>0$, for every $\widehat{\varepsilon}_1,\widehat{\varepsilon}_2>0$, we can take $\tau\le\widehat{\delta}_1(\widehat{\varepsilon}_1)$ and $\tau'\le\widehat{\delta}_2(\widehat{\varepsilon}_2)$ and there will exist $\widetilde{\theta}_i,\overline{\theta}_i\in\Theta_i$ with the above properties.

Finally, we can apply *Lemma 75* for $\theta_{-i}\in\prod_{j\neq i}\Theta_j$ , for $\widehat{g}=g$ and for the above $\widetilde{\theta}_i,\overline{\theta}_i\in\Theta_i$ to get: $\forall\phi>0$, $\exists\rho>0$ such that $g(\theta_i,\theta_{-i})\in B_\rho(\widehat{g}(\theta_i',\theta_{-i}))\Rightarrow\ q_i(\theta_i,\theta_{-i})\in B_\phi(q_i(\theta_i',\theta_{-i}))$, whenever we choose $\delta,\widehat{\varepsilon}_1,\widehat{\varepsilon}_2>0$ such that, for any given $\phi>0$, $\rho(\phi)\ge\delta+\tau+\tau'$. Then, by (i) and (ii), it holds: $\exists\widetilde{\theta}_i,\overline{\theta}_i\in\Theta_i$ such that $g(\overline{\theta}_i,\theta_{-i})\in B_\rho(g(\widetilde{\theta}_i,\theta_{-i}))$, so $\forall\phi,\delta,\widehat{\varepsilon}_1,\widehat{\varepsilon}_2>0$, $\exists\widetilde{\theta}_i,\overline{\theta}_i\in\Theta_i$ such that:

(1). $q_i(\overline{\theta}_i,\theta_{-i})\in B_\phi(q_i(\widetilde{\theta}_i,\theta_{-i}))$,

(2). $q_i(\overline{\theta}_i,\theta_{-i})\in B_{\widehat{\varepsilon}_1}(q_i(\lim_{\tau\to 0}\widehat{\theta}_i(\tau,\overline{k}(\delta),\theta_{-i}),\theta_{-i}))$ and

(3). $q_i(\widetilde{\theta}_i,\theta_{-i})\in B_{\widehat{\varepsilon}_2}(q_i(\lim_{\tau'\to 0}\widehat{\theta}_i(\tau',k,\theta_{-i}),\theta_{-i}))$.

Now, choose $\phi,\widehat{\varepsilon}_1,\widehat{\varepsilon}_2>0$ sufficiently small such that, for example, $\phi+\widehat{\varepsilon}_1+\widehat{\varepsilon}_2\le\dfrac{\varepsilon}{4}$, to observe that (1), (2) and (3) makes any $\varepsilon>0$ such that $\forall\delta>0,\ \exists\overline{k}(\delta)\in B_\delta(k),\ \widehat{q}_i'(\overline{k},\theta_{-i})\notin B_\varepsilon(\widehat{q}_i'(k,\theta_{-i}))$ impossible, so we enter into a contradiction and the function $\widehat{q}_i'(k,\theta_{-i})$ is continuous for all $k\in K$.

Now we will consider the same kind of composed function of *Theorem 58*: $\widehat{P}_i:K\times\prod_{i=1}^n\Theta_i\to E$ defined as: $\forall k\in K,\ \forall\theta_i\in\Theta_i,\ \forall\theta_{-i}\in\prod_{j\neq i}\Theta_j$,

$\widehat{P}_i(k,\theta_i,\theta_{-i})=P_i(v_i(k,\theta_i),\widehat{q}_i(k,\theta_{-i}))$. We know now that, given any $\theta_{-i}\in\prod_{j\neq i}\Theta_j$, this function is continuous in the whole $K$, since $\Theta_i$ are continuous by hypothesis and we have proved yet that $\widehat{q}_i$ are continuous functions. Hence suppose, by contradiction, that the Theorem is not true: $\exists i\in N,\ \exists\overline{\theta}_{-i}\in\prod_{j\neq i}\Theta_j,\ \exists\theta_i\in\Theta_i,\ \exists k\in K$ such that: $\widehat{P}_i(g(\theta_i,\overline{\theta}_{-i}),\theta_i,\overline{\theta}_{-i})<\widehat{P}_i(k,\theta_i,\overline{\theta}_{-i})$ for some selection $g$ from $K^*$-note that in other case, $\widehat{P}_i\in W(K^*)$-, or in other words, the following statements are true:

(i). $\exists\widehat{\varepsilon}>0$ such that $\widehat{P}_i(k,\theta_i,\overline{\theta}_{-i})-\widehat{P}_i(g(\theta_i,\overline{\theta}_{-i}),\theta_i,\overline{\theta}_{-i})=\widehat{\varepsilon}$.

(ii). Since $K^*$ is individually quasi-decisive, for $k,\overline{\theta}_{-i}$ and any $\tau>0$, there exist a selection from $\widehat{\theta}_i$ such that $\exists\overline{k}\in K$ such that $\overline{k}\in K^*(\widehat{\theta}_i(\tau,\overline{k},\overline{\theta}_{-i}),\theta_{-i})\in K\ \&\ \ k\in B_\tau(\overline{k})$.

(iii). Since given $\theta_i$ and $\overline{\theta}_{-i}$, $\widehat{P}_i$ is continuous in $K$, in particular for $k$ it is true that: $\forall\varepsilon > 0$, $\exists\delta(k,\varepsilon) > 0 \ni \underline{k} \in B_\delta(k) \Rightarrow P_i(\underline{k},\theta_i,\overline{\theta}_{-i}) \in B_\varepsilon(P_i(k,\theta_i,\overline{\theta}_{-i}))$.

First, using (iii), for $k \in K$, for $\widehat{\varepsilon} > 0$, $\exists\delta(k,\widehat{\varepsilon}) > 0 \ni \widetilde{k} \in B_{\delta(k,\widehat{\varepsilon})}(k)$ -by (b)- $\Rightarrow \widehat{P}_i(\widetilde{k},\theta_i,\overline{\theta}_{-i}) \in B_{\widehat{\varepsilon}}(\widehat{P}_i(k,\theta_i,\overline{\theta}_{-i}))$. Now, let us set $\tau = \delta(k,\widehat{\varepsilon})$; by (ii), for this $\tau$ there will exist $\widetilde{\theta}_i \ni \widetilde{k} \in K^*(\widetilde{\theta}_i(\delta(k,\widehat{\varepsilon}),k,\overline{\theta}_{-i}),\overline{\theta}_{-i}) \in K$ & $\widetilde{k} \in B_{\delta(k,\widehat{\varepsilon})}(k)$. Finally, this last expression can be written as: for $k \in K$ and given $\theta_i$ and $\overline{\theta}_{-i}$, $\exists\widetilde{\theta}_i(\delta(k,\widehat{\varepsilon}),k,\overline{\theta}_{-i})$, $\exists\widetilde{k} \in K^*(\widetilde{\theta}_i,\overline{\theta}_{-i}) \in K$ s.t.

$\left|\widehat{P}_i(k,\theta_i,\overline{\theta}_{-i})) - \widehat{P}_i(\widetilde{k},\theta_i,\overline{\theta}_{-i})\right| < \widehat{\varepsilon}$ and $\widehat{P}_i(k,\theta_i,\overline{\theta}_{-i}))-\widehat{P}_i(g(\theta_i,\overline{\theta}_{-i}),\theta_i,\overline{\theta}_{-i}) = \widehat{\varepsilon}$ by definition, so it has to be that for $k \in K$, $\theta_i \in \Theta_i$ and $\overline{\theta}_{-i} \in \prod_{j\neq i}\Theta_j$,

$\exists\widetilde{\theta}_i(\delta(k,\widehat{\varepsilon}),k,\overline{\theta}_{-i})$, $\exists\widetilde{k} \in K^*(\widetilde{\theta}_i,\overline{\theta}_{-i}) \in K$ such that there exist the following selection from $K^*$ :

$\widetilde{g}(\theta) = \begin{cases} \widetilde{k} & iff\ (\theta_i,\theta_{-i}) = (\widetilde{\theta}_i,\overline{\theta}_{-i}) \\ g(\theta_i,\theta_{-i}) & otherwise \end{cases}$ such that

$\widehat{P}_i(\widetilde{g}(\widetilde{\theta}_i,\overline{\theta}_{-i}),\theta_i,\overline{\theta}_{-i}) > \widehat{P}_i(\widetilde{g}(\theta_i,\overline{\theta}_{-i}),\theta_i,\overline{\theta}_{-i})$, which is the same that

$$P_i(v_i(\widetilde{g}(\widetilde{\theta}_i,\overline{\theta}_{-i}),\theta_i),q_i(\widetilde{\theta}_i,\overline{\theta}_i)) > P_i(v_i(\widetilde{g}(\theta_i,\overline{\theta}_{-i}),\theta_i),q_i(\theta_i,\overline{\theta}_{-i}))$$

so the compensation mechanism fails to be incentive compatible for some $i$: a contradiction.

$\Leftarrow$) The sufficiency condition is exactly the same that the one in *Theorem 58.* ■

*Theorem 76* shows that our main result is robust even if we enlarge the set of admissible social rules to individually quasi-decisive SCCs. The price to pay is assuming the continuity of the compensation mechanisms, a property that does not seem extremely restrictive when working with our complex set of feasible alternatives. Nevertheless, we can say nothing about the set of discontinuous mechanisms implementing individually quasi-decisive SCCs.

# References


[1] Barberà, S. & Peleg, B.:"Strategy-Proof Voting Schemes with Continuous Preferences," Social Choice and Welfare, 7, (1990), 31-38.

[2] Clarke, E.H.: "Multipart Pricing of Public Goods," Public Choice, 8 (1971), 19-33.

[3] Dasgupta, P.,Hammond, P. & Maskin, E.:"The Implementation of Social Choice Rules," Review of Economic Studies, 46, (1979), 185-216.

[4] Gibbard, A.:"Manipulation of Voting Schemes: A General Result," Econometrica, 41, (1973), 587-601.

[5] Green, J. & Laffont, J.J..:"Characterization of Satisfactory Mechanism for the Revelation of Preferences for Public Goods," Econometrica, 45, (1977), 427-438.

[6] ———:"Incentives in Public Decision Making," (1979). Amsterdam: North-Holland.

[7] Groves, T.:"Incentives in Teams," Econometrica, 41 (1973), 617-631.

[8] ———:"Efficient Collective Choice when Compensation is Possible," Review of Economic Studies, (1979), 227-241.

[9] Groves, T. & Loeb, M.:"Incentives and Public Inputs," Journal of Public Economics, 4, (1975), 211-226.

[10] Hurwicz, L.:"On Informationally Decentralized Systems," in Decision and Organization, ed. by B. McGuire and R. Radner. Amsterdam: North-Holland, (1972), 297-336.

[11] Hurwicz, L. & Walker, M.:"On the Generic Nonoptimality of Dominant-Strategy Allocation Mechanisms: A General Theorem that Includes Pure Exchange Economies," Econometrica, 58, (1990), (683-704).

[12] Moulin, H.:"On Strategy-proofness and Single Peakedness," Public Choice, 35, (1980), 437-455.

[13] Roberts, K.:"The Characterization of Implementable Choice Rules," in "Aggregation and Revelation of Preferences," edited by J.J. Laffont. North-Holland Publishing Company, (1979).

[14] Satterthwaite, M. A.:"Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," Journal of Economic theory, 10, (1975), 187-217.

[15] Walker, M.:"On the Nonexistence of a Dominant Strategy Mechanism for Making Optimal Public Decisions," Econometrica, 48, (1980), 1521-1540.

# 5 DOMINANT STRATEGIES IMPLEMENTATION OF THE CRITICAL PATH ALLOCATION IN THE PROJECT PLANNING PROBLEM

## 5.1 Introduction

In this chapter we study the existence of dominant strategy mechanisms in an incomplete information version of the *Critical Path Method* (CPM) or *Program Evaluation and Review Technique* (PERT), a well-known solution to a standard production network problem in the Operations Research literature -see, for example, Bazawa and Jarvis [1], Derigs [2] and Deo and Pang [3] as comprehensive introductions to the topic.

The PERT was first developed and used, with great success, during the late 1950s by the US Navy to control the progress of the construction of the Polaris missiles, an extraordinary complex project carried out by different production units. The CPM technique was found independently and applied to virtually the same kind of problems, although the PERT was a little bit more general since it allowed for some degree of uncertainty.

The simplest *production network problem* at which these techniques are applied is very simple: A project consisting in a number of different elementary tasks has to be carried out. Each task constitutes a time-consuming production activity -abstracting from any other economic resource employed- necessary for the completion of the project. Tasks cannot generally be allocated arbitrarily, since a particular task may need some others to be finished before it starts, and maybe some of its preceding tasks are also preceded by others, and this network structure is what generates the complexity of the problem. Of course, for the problem to make sense some technological restrictions must be introduced: cycles and loops of precedence are technologically unfeasible. The CPM and the PERT are equivalent methods to analyze the sequences of tasks such that the total amount of time needed to finish all the tasks is minimized[18]. A *critical path* is a sequence of tasks

[18] There is an underlying assumption that the total cost of the completion of the project to the planner or principal is increasing in the total amount of time needed.

that are undertaken one after the other and such that the completion of the sequence requires exactly the minimum amount of time needed to terminate all tasks. The CPM and the PERT define algorithms to identify the critical paths and those tasks with some roominess -the tasks outside the critical paths that can therefore be allocated at different starting times without affecting the total (minimum) duration of the project-. It can be proven that the whole allocation problem can be transformed into a linear programming one and the operations research analysis is essentially one of computability. The CPM and the PERT also assume that the planner knows all relevant data about the technologies: tasks, time needed by each one and which are the immediate preceding tasks of each. The PERT admits some uncertainty about the duration of each task, but the probability distribution is always known to the planner.

This paper explores the same task allocation problem but assumes that each task is carried out by an economic agent[19]; as in many production problems, the agent responsible for each task can be a worker, supervisor, firm or the people in charge of a sub-project or division within a firm, but the key assumption is that she is better informed about the technological characteristics of her particular task than the planner herself. In the limit, we assume that the planner has to rely on the information reported by the agents in each task to allocate the times and sequence of the tasks and the rest of the agents do not need to know anything about a different task but their own. The agents are rational and will exploit their informational advantage if given the opportunity. Nevertheless, the planner is not completely uninformed: if some agents lie about the duration or precedence of their tasks and they are allocated in an unfeasible way, they will be caught and punished hard enough to discourage any agent to lie in a potentially detectable way. Notice, however, that there is still room enough for safe lies: reporting longer duration times and declaring as preceding tasks more tasks that those really needed are always undetectable when the planner is allocating tasks by using the CPM or the PERT methods, and they are actually the only lies that we will allow in this paper. We also assume that the agents, if not compensated in other way, are interested in delaying the completion of their own task as much as possible. This is justified because each hour employed

[19]If the tasks are undertaken by machines or automata alone, the original task-machine allocation problem still applies, but it seems unlikely that no worker (but the planner herself) is involved in any task in real life problems.

in carrying out the task requires a costly effort to be made by the agent and agents prefer the same disutility to take place later than sooner. There is, therefore, a fundamental conflict of interests between the planner or principal -who is using the PERT with the information reported by the agents and who tries to minimize the total cost of the project- and her agents, who would prefer the project to be delayed forever. A simplifying and important assumption we impose is that disutility of effort is known and the same for everybody and is normalized to unity, but we still allow for differences in the agents' relative impatience -their time discounts-. Of course, planning the production network and the starting-finishing time of each task is not the only way the planner has to influence the behavior of the agents. We assume that the planner can design a *transfer scheme* depending on the reported technologies that specifies the monetary payments that each agent receives or has to pay -taxes- to the planner[20]. This rule is known by the agents, who have committed to work in the project and cannot quit if they think that they will not receive enough money and money enters additively on the agents' payoffs. Our main aim is to analyze the possibility of designing *anonymous, strategy-proof transfer schemes*, i.e., a payoff structure such that no agent has an incentive to lie about her technology regardless of what other agents report to the planner and whatever her own technology is. We then examine two additional constraints imposed on the payments structure: first, we study the possibility of designing *individually rational payments* and then we assume that the planner has agreed -or is payed- a total price for the completion of the project, which is given in the problem and constitutes a budget constraint for her. This budget has to be completely shared by the agents, so that we are imposing a *balance* condition.

The chapter proceeds as follows: In *Section 2* the formal model is introduced and the definitions properly stated. *Section 3* deals with the results and we conclude with some comments.

## 5.2 The model

Let $\overline{N}$ be a potential finite set of productive agents and $N = \{1, ..., n\} \subseteq \overline{N}$ be a subset of *agents* indexed by $i, j, k, l, z \in N$. The total number of agents in set $N$ is $n \geq 2$. Each agent has to perform a *task* for the completion of

[20] We assume that the transfers are implemented before the task allocation starts. Therefore, in period "0" the agents report their technologies and the transfers are made.

a *project*. We do not allow in this model for multi-task agents to simplify the problem and we are not concerned with the matching allocation problem between agents and tasks. Hence, we assume that each agent either is the only -or the best- agent capable to perform a particular task or that the matching or allocation of tasks to agents took place in an earlier stage and is given. Therefore, there is no reason to define a separate set of tasks and we identify the set of tasks with the set of agents $N$. Each task -or alternatively from now on, each agent- is characterized by its belonging to a network such that a given task cannot be undertaken before some other tasks are finished. Moreover, carrying out each task is a time-consuming process and some work effort -or maybe some cost of capital- has to be invested in order to be successfully completed. Given any task $i \in N$, we denote as $P_i \in 2^N$ the set of *preceding tasks* of task $i$. For the problem to make sense, we need to impose some minimal structure to the admissible sets of preceding tasks for all tasks. In particular, the following constraints should hold in any well-defined problem: we say that a project is *technologically feasible* if the following conditions hold:

(i). Temporal irreflexivity: $\forall i \in N,\ i \notin P_i$.

(ii). Temporal asymmetry: $\forall i, j \in N,\ i \neq j,\ i \in P_j \longrightarrow j \notin P_i$.

(iii). Temporal transitivity: $\forall i, j, k \in N,\ i \neq j \neq k,\ i \in P_j \ \&\ j \in P_k \longrightarrow k \notin P_i$.

Condition (i) establish that no task is ever preceded by itself. Condition (ii) prevents two different tasks to precede each other and (iii) rules out cycles of precedence. Given the linear structure of time, the meaning of the above properties become obvious.

Furthermore, the technology requires the use of costly time for undertaking each task. For simplicity, we assume time to be discrete -measured in any relevant unit-. Henceforth, time intervals might be hours, minutes or days, but let us call them hours-. Let $E$ be the real line and $Z_+$ be the set of non-negative integers -time structure considering that 0 stands for *now*- and $Z_{++}$ be the set of positive integers. $T_i \in Z_{++}$ is the minimum number of hours that task $i$ needs to be terminated, given the optimal use of the resources available and given that preceding tasks -and preceding tasks of its preceding tasks and so on- have been done before. We assume that each hour employed by the agent to the completion of his task entails a disutility of effort -or some depreciation of the use of capital-. Moreover, agents discount future effort with respect to effort now at time 0 -when the allocation has to be decided-, but they will not have any cost until their own task has to be performed. For instance, if agent $i$´s task lasts -efficiently done- $T_i$ hours and it is allocated to start at time $t_i^0 \in Z_+$, the disutility of the agent performing the task is -measured in some monetary unit- $-\sum_{t=t_i^0}^{t=t_i^0+T_i-1} \beta_i^t$ , where the time discount applied by each agent ($\beta_i$, with $0 < \beta_i < 1$) is a function of the intrinsic difficulty of the task and of the laziness of the workers, which is part of their private information.[21] We further assume that disutility of effort done in period 0 (now) of any agent is known to the planner and equal to 1[22]. We call a *project planning economy* to a tuple of the form: $e = \langle N, P_i, T_i, \beta_i \, \forall i \rangle = \{N, e_1, ..., e_i, ..., e_n\}$ , provided that the project is technologically feasible[23]. Let $PPE$ be the set of all project planning economies. Given an economy $e \in PPE$, a feasible allocation -or simply, an *allocation*-, denoted by $x, y, z \in Z_+^{2(\#N)}$ is a vector that assigns a pair of *times* to each task or agent, $x = (t_1^0, t_1^1, t_2^0, t_2^1, ..., t_i^0, t_i^1, ..., t_n^0, t_n^1)$ with the following properties:

(a). $\exists i \in N$ such that $t_i^0 = 0$.

(b). $\forall i \in N,\ t_i^1 - t_i^0 \geq T_i$.

(c). $\forall i, j \in N,\ i \in P_j \longrightarrow t_j^0 \geq t_i^1$.

An allocation establishes a technologically feasible plan for the tasks to be carried out on time: $t_i^0$ stands for the planned starting time of task $i$ and $t_i^1$ denotes the date after the termination date of task $i$. (a) means that some task should be initiated in period 0 (when the allocation is decided), (b) establish that no task should be allocated a working time smaller than the minimum time required to be done and (c) requires that no task can be started before all its preceding tasks have been completed. Let $FA(e)$ be the set of feasible allocations for economy $e$. Given an economy $e \in PPE$, a *critical path allocation* ($CPA$) is an allocation $\overline{x} = (\overline{t}_1^0, \overline{t}_1^1, ..., \overline{t}_i^0, \overline{t}_i^1, ..., \overline{t}_n^0, \overline{t}_n^1) \in FA(e)$ such that $\max_{i \in N} \overline{t}_i^1 \leq \max_{i \in N} t_i^1 \ \forall x = (t_1^0, t_1^1, ..., t_i^0, t_i^1, ..., t_n^0, t_n^1) \in$

[21] Note that the discount used by task $i$, $\beta_i$ may also be interpreted as the cost -depreciation, funding cost,...- of the capital used in that task.

[22] This is partly a simplifying assumption. For our results to hold it is needed that the planner either know the (possibly different) agents' disutility of effort or at least know that there is some common upper bound on every agent disutility of effort, which does not seem to be an unreasonable assumption. In that case, the mechanism we propose should be properly re-scaled.

[23] Our definition of an economy includes the set of agents, that is allowed to vary within the range of admissible economies. Mechanisms that work for some set of $PPE$ have to take into account that the number of agents could be different. The identity of agents and tasks makes the problems consistent.

$FA(e)$, i.e., an allocation such that the period of time until the last task is finished is as small as possible within all the feasible allocations. Let us denote by $CPA(e)$ the set of $CPAs$ for an economy $e$. An *efficient* $CPA$ or $CPA^{+}(e)$ is a critical path allocation such that no agent can be better off in any other critical path, i.e., $\overline{x} = (\overline{t}_1^0, \overline{t}_1^1, ..., \overline{t}_i^0, \overline{t}_i^1, ..., \overline{t}_n^0, \overline{t}_n^1) \in CPA(e)$ is efficient[24] iff $-\sum_{t=\overline{t}_i^1-T_i}^{t=\overline{t}_i^1-1} \beta_i^t \geq -\sum_{t=t_i^1-T_i}^{t=t_i^1-1} \beta_i^t \;\forall\, x = (t_1^0, t_1^1, ..., t_i^0, t_i^1, ..., t_n^0, t_n^1) \in CPA(e)$. Notice that under our assumptions, $CPA^{+}(e)$ is a singleton for each economy.

There are several ways to find the efficient $CPA$ for a given economy. The optimization problem can be formulated as one of linear programming or different algorithms can be applied to reach the solution. In what follows, we will use the following *strings* $CPA^{+}$ *algorithm:* given any economy $e$, the set of $CPA^{+}(e)$ comes from following the steps:

*Step 1*: Take tasks 1 to $n$. Assign the negative integers $z_i^0$ and $z_i^1 \in Z_{-}$: $z_i^1 = 0, z_i^0 = -T_i$.

*Step 2*: Take task $i = 1, ...n$ successively. In each sub-step $i$, do: $\forall j \in P_i$, redefine $z_j^1 = z_i^0$ and $z_j^0 = z_j^1 - T_i$.

*Step 3*: Repeat Step 2 until no new change emerges.

*Step 4*: $\forall i \in N$, redefine $\overline{t}_i^0 = z_i^0 + \min_{i\in N} z_i^0$ and $\overline{t}_i^1 = z_i^1 + \min_{i\in N} z_i^0$. $\overline{x} = (\overline{t}_1^0, \overline{t}_1^1, ..., \overline{t}_i^0, \overline{t}_i^1, ..., \overline{t}_n^0, \overline{t}_n^1) = CPA^{+}(e)$.

REPEAT STEPS 1 TO 4 UNTIL NO FURTHER CHANGE OCCURS.

STOP.

Given any economy $e \in PPE$, we will be interested in those tasks that are critical. We call a *critical string associated to the* $CPA(e)$ to a sequence of subsets of tasks $\left\{S^1(e), S^2(e), ..., S^k(e)\right\}$, with $S^h(e) \subseteq N$ and $S^h(e) \cap S^l(e) = \emptyset \;\forall h, l$ that can be found with the following algorithm using the $CPA^{+}(e)$ allocation:

*Step 1*: Take any $i \in N$ such that $\overline{t}_i^0 = 0$. $i$ belongs to the first set of tasks in the sequence: $i \in S^1$.

*Step 2:* Take any $j \in N$ such that $\overline{t}_j^0 = \overline{t}_i^1$ for any $i \in S^1$ and $i \in P_j$. $j$ belongs to the second: $j \in S^2$.

....

*Step k*: Take any $i \in N$ such that $\overline{t}_i^0 = \overline{t}_j^1$ for any $j \in S^{k-1}$ and $j \in P_i$. $i$ belongs to the $k$th set (we assume that there are $k(e) \in \{1, ..., n\}$

[24] We assume that every task is efficiently performed and whenever more time than needed is provided in the allocation, only the last $T_i$ hours will be used, and hence achieving the smaller possible disutility.

subsets). Let us call $S(e)$ the union of the tasks belonging to any critical string associated to the $CPA^{+}(e)$, i.e., $S(e) = \bigcup_{h=1}^{k} S^{h}(e)$.

Notice that under our assumptions both algorithms work in selecting the unique $CPA^{+}(e)$ and the critical strings that define the minimum period of time needed for all the tasks -the project- to be finished. Moreover, since it always hold that $\forall e \in PPE,\ \forall h \in \{1, ..., k(e)\}, \forall i, j \in S^{h}(e),\ T_i = T_j$, we denote by $T_{S^h} = T_i = T_j$ to the generic duration of every task in $S^h$ (omitting the reference to a given $e$ for simplicity). In what follows, we focus on implementing the $CPA^{+}(e)$. This is justified because that selection is the only one that maximizes the total welfare of the agents subject to the planner achieving her total time-minimising objective -assuming that the planner receives a payoff strictly increasing in the total duration of the project-. Moreover, other selections from the $CPA(e)$ like that which allocates tasks not belonging to the critical path to start as early as possible are even more difficult to implement because that rules defy more directly the agents' incentives by imposing an inefficient cost on them -starting early-.

A *Project Planning function (PPF)* is a function that assigns a feasible alternative to every admissible economy, i.e., $\varphi : PPE \longrightarrow FA(e)$. We say that a $PPF$ is a *critical path PPF* if and only if $\forall e \in PPE,\ \varphi(e) \in CPA(e)$. An *efficient critical path PPF* is a $PPF$ such that $\forall e \in PPE,\ \varphi(e) = CPA^{+}(e)$. Given an economy $e \in PPE$ and a $PPF$ $\varphi$, $\forall i \in N,\ \varphi_i^0(e)$ will denote the component function relative to the starting time of agent $i's$ task and $\varphi_i^1(e)$ denotes the component function giving the allocation of the end of agent $i$'s task.

The overall interest of the organization is modeled as the objectives of the *planner or principal.* We are assuming all along the paper, following the traditional PERT literature, that the cost of the project to the principal is proportional to the maximum amount of time spent for its completion. Let us assume that the planner wants to minimize the length of the project by selecting always $CPA$ allocations. If the planner is perfectly informed about the relevant economy $e$, it should not be difficult for him to apply the PERT techniques or the linear programming version of them to find the $CPA$ allocations. But we are not concerned in this paper about how to find this allocations, but about the possibility for the planner to achieve those outcomes when she is not informed about the technologies. We assume that each agent is better informed about the characteristics of his own task than the planner -or even any other agent-, so both the minimum duration of the task, $T_i$, the set of preceding tasks, $P_i$, and the time discount $\beta_i$ are agent $i$´s private information ($e_i = (T_i, P_i, \beta_i)$). The planner can only decide the final allocation based on the *reported technologies*, denoted as $\widehat{e}_i = (\widehat{T}_i, \widehat{P}_i, \widehat{\beta}_i)$ of each agent. Therefore, the planner is interested in designing a *direct revelation mechanism*[25] such that the agents will have no incentive to lie about their true technologies. However, the planner still knows some information about the relevant set of agents or tasks involved in the project, $N$, and the consistency of the whole project -$e \in PPE$-. Hence, we will assume that, given an economy $e$, the final allocation given the agents' reported technologies has to be technologically feasible, i.e., the planner can always find a lie if when the project is not technologically feasible, some task cannot be undertaken by an agent given his reported $\widehat{e}_i$. We assume that any detected lie can be heavily punished in such a way that no agent is ever interested in reporting a set $\widehat{P}_i \subset P_i$.[26]Using an identical reasoning, no agent can ever use a lie such that $\widehat{T}_i < T_i$.[27] Notice that the agents can still lie by using $\widehat{P}_i \supset P_i$ and $\widehat{T}_i > T_i$ -trying to delay the completion of the project in order to avoid early costs- if they are not given other additional incentives. We allow for monetary transfers to the agents based on the agents' reported technologies, but we assume that the total amount to be transferred to the agents is a fixed quantity -the price of the project-. Now, we define the concept of an incentive compatible mechanism in this setting. Given two economies $e = \langle N, P_i, T_i, \beta_i \; \forall i \rangle = \{N, e_1, ..., e_i, ..., e_n\}$ and $e' = \langle N, P'_i, T'_i, \beta'_i \; \forall i \rangle = \{N, e'_1, ..., e'_i, ..., e'_n\}$, we write $e'_i \subset e_i$ whenever $P'_i \subseteq P_i$ and $T'_i \subseteq T_i$ and $e' \subset e$ when $e'_i \subset e_i \; \forall i \in N$. We shall also make use of the following well-known notation to avoid large expressions: $e = (N, e) = (N, e_S, e_{-S})$, $\forall S \subseteq N$, and in particular, for $S = \{i\}$, $e = (N, e_i, e_{-i})$. [28]

**Definition 77** *A* ***mechanism*** *$M$ is a set of transfer functions $\{w_i \; \forall i \in N\}$ of the kind $w_i : PPE \longrightarrow E$ for every set of agents $N \subseteq \overline{N}$.*

Notice that the above definition entails that mechanisms are direct: the only information used by the planner to allocate transfers are the agents'

[25] A direct revelation mechanism asks the agents about their *types*.

[26] Notice that actually any agent that reports a narrower set of preceding tasks, either leads to her detection or cannot change the allocation, so it is not individually rational to do so and we can therefore eliminate all those irrelevant strategies.

[27] Basically, we are assuming that the planner can monitor when the agents start their tasks and when they finish.

[28] This definitions hold irrespective of $\beta_i \; \forall i \in N$.

revealed technologies -their types-.

**Definition 78** *A mechanism* $M = \{w_i \; \forall i \in N\}$ ***implements an efficient critical path*** $PPF$ $\varphi$ *if the following holds:*
$\forall e \in PPE$, $\forall i \in N$, $\forall e'_i \supset e_i$,

$$w_i(e) - \sum\nolimits_{t=\varphi_i^1(e)-T_i}^{t=\varphi_i^1(e)-1} \beta_i^t \geq w_i(N, e'_i, e_{-i}) - \sum\nolimits_{t=\varphi_i^1(N,e'_i,e_{-i})-T_i}^{t=\varphi_i^1(N,e'_i,e_{-i})-1} \beta_i^t$$

*We also say that these mechanisms are* ***strategy-proof****.*

A mechanism implements the efficient critical path $PPF$ iff any agent, by reporting a different technology cannot improve her net payoff -the transfer received minus the disutility of effort-, and this whatever her true technology is and regardless the others' reported technologies. Therefore, we are interested in a strong incentive compatibility property.

**Definition 79** *Given any positive number* $C > 0$*, a mechanism* $M^C = \{w_i \; \forall i \in N\}$ *is* ***balanced*** *if* $\forall e \in PPE$, $\sum_{i=1}^n w_i(e) = C$.

This property imposes that the transfer or reward scheme designed by the planner has to be balanced -the whole budget coming from the project should be distributed among the agents that are involved in it-.

**Definition 80** *A balanced mechanism* $M^C = \{w_i \; \forall i \in N\}$ *is* ***invariant to the project size*** *if* $\forall e = (N, P_i, T_i, \beta_i) \in PPE, \forall \lambda > 0,$ *if* $e' = (N, P_i, \lambda T_i, \beta_i)$, $M^{\lambda C} = \{w_i(e') \; \forall i \in N\} = \{\lambda w_i(e) \; \forall i \in N\}$.

A balanced mechanism is invariant to the project size if the transfers are proportionally affected by a proportional re-scaling of the project; for example, doubling the tasks minimum durations joint with the project value $C$ should double every agent transfers. This property may be desirable because it introduces some fairness criterion to the sharing rule when the project is re-scaled: the agent's relative payoffs do not change when we measure the resource "time" in hours, minutes, days or months.

**Definition 81** *Given any reservation utility* $\overline{U} \in Z$*, a mechanism* $M = \{w_i \; \forall i \in N\}$ *implementing* $PPF$ $\varphi$ *is* ***individually rational*** *if*
$\forall e \in PPE$, $\forall i \in N$,

$$w_i(e) - \sum\nolimits_{t=\varphi_i^1(e)-T_i}^{t=\varphi_i^1(e)-1} \beta_i^t \geq \overline{U}$$

Individual rationality requires the payoffs to be designed such that for every economy no agent gets a too small payoff that might lead the agent to leave the project is possible. An implicit simplifying assumption is that a net utility of $\overline{U}$ constitutes the common agents' reservation utility threshold for project acceptance.

**Definition 82** *Given a PPF $\varphi$, a mechanism $M = \{w_i\}$ is* ***innovation-monotonic*** *if the following holds:*
$\forall e \in PPE$, $\forall i \in N$, $\forall e_i' \supset e_i$,

$$w_i(e) - \sum_{t=\varphi_i^1(e)-T_i}^{t=\varphi_i^1(e)-1} \beta_i^t \geq w_i(N, e_i', e_{-i}) - \sum_{t=\varphi_i^1(N,e_i',e_{-i})-T_i'}^{t=\varphi_i^1(N,e_i',e_{-i})-1} \beta_i^t$$

A mechanism is innovation-monotonic iff an innovation that makes any agent more efficient cannot make him be worse off. Notice that every mechanism that implements the critical path $PPF$ has to be innovation monotonic given the efficient critical path $PPF$. This property constitutes an additional justification for both implementing the efficient critical path $PPF$ and using strategy-proof mechanisms.

**Definition 83** *A $PPF$ $\varphi$ is* ***anonymous*** *if for all $N \subseteq \overline{N}$, $e \in PPE$ and any permutation $\sigma(N)$ of the agents, the following holds: $\varphi_i^k(N, e_i, e_{-i}) = \varphi_{\sigma(i)}^k(N, e_{\sigma(i)}, e_{-\sigma(i)})$, for $k = \{0, 1\}$ and for all $i \in N$.*

This requirement is an obvious fairness property that excludes $PPFs$ that take into account the agents' names and not only their technology. Notice that the efficient critical path $PPF$ are anonymous.

**Definition 84** *A mechanism $M = \{w_i\}$ is* ***anonymous*** *if for all $N \subseteq \overline{N}$, $e \in PPE$ and any permutation $\sigma(N)$ of the agents, the following holds: $w_i(N, e_i, e_{-i}) = w_{\sigma(i)}(N, e_{\sigma(i)}, e_{-\sigma(i)})$ for all $i \in N$.*

Again, anonymity establish that the information about the agents' names is not used to allocate the transfers.

Now, we will proceed with the main results in the paper.

## 5.3 Results

Our first possibility result proves the existence of anonymous and balanced mechanisms implementing the efficient critical path $PPF$. To prove the theorem, we still need some definitions.

Let us call a *string* associated to agent $i \in N$ and economy $\widehat{e} \in PPE$, denoted as $R^i(\widehat{e})$, to the set of agents obtained with the following algorithm:

Step 1: take a single $j \in \widehat{P}_i$. $j \in R^i(\widehat{e})$. If there is no such an agent, $R^i(\widehat{e}) = \emptyset$ and the process stops.

Step 2: take a single $k \in \widehat{P}_j$. $k \in R^i(\widehat{e})$. If there is no such an agent, the process stops.

Step 3: take any $h \in P_k$. $h \in R^i(\widehat{e})$. If there is no such an agent, the process stops.

........

Eventually, since $N$ is finite and technologies are feasible, the algorithm stops and it is clear that any string $R^i(\widehat{e}) \subseteq N \setminus \{i\}$. Now, take the union of all the strings associated to agent $i \in N$ and declared technology $\widehat{e} \in PPE$, i.e., the set of agents with tasks that should necessarily precede $i$. Let us denote as $r^i(\widehat{e})$ the total number of agents belonging to any string associated to agent $i \in N$ and technology $\widehat{e} \in PPE$, i.e., $r^i(\widehat{e}) = \# \bigcup R^i(\widehat{e})\ \forall i \in N,\ \forall \widehat{e} \in PPE$. Notice that $r^i(\widehat{e}) = 0$ if and only if $\widehat{P}_i = \emptyset$. Let us call *terminal agents* to the set of agents such that have not declared to have any preceding task and let us denote them as $S(\widehat{e})$, i.e., $\forall \widehat{e} \in PPE,\ S(\widehat{e}) = \left\{ i \in N \text{ s.t. } \widehat{P}_i = \emptyset \right\}$, or identically, $S(\widehat{e}) = \{i \in N \text{ s.t. } r^i(\widehat{e}) = 0\}$. Notice also that there exist always at least one terminal agent for every feasible technology. Now, we state our main result in this chapter:

**Theorem 85** *There exist anonymous, balanced and invariant to the project size mechanisms implementing the efficient critical path PPF.*

**Proof.** Let us consider the following mechanism: $\forall \widehat{e} \in PPE$, $\forall N \in \overline{N}$, $\forall i \in N$, $w_i(N, \widehat{e}_1, ..., \widehat{e}_n) = \dfrac{C}{n} +$

$$\begin{cases} -(r^i(\widehat{e})+1)\widehat{T}_i & if\ i \notin S(\widehat{e})\ and\ \#S(\widehat{e})>1 \\ \dfrac{\sum_{j\notin S(\widehat{e})}(r^j(\widehat{e})+1)\widehat{T}_j}{\#S(\widehat{e})} - \widehat{T}_i + \dfrac{\sum_{j\in S(\widehat{e})/\{i\}}\widehat{T}_j}{\#S(\widehat{e})-1} & if\ i \in S(\widehat{e})\ and\ \#S(\widehat{e})>1 \\ -(r^i(\widehat{e})+1)\widehat{T}_i + \dfrac{\sum_{j\in S(\widehat{e})}\widehat{T}_j}{n-1} & if\ i \notin S(\widehat{e})\ and\ \#S(\widehat{e})=1 \\ \sum_{j\notin S(\widehat{e})}(r^j(\widehat{e})+1)\widehat{T}_j - \widehat{T}_i & if\ i \in S(\widehat{e})\ and\ \#S(\widehat{e})=1 \end{cases}$$

In words, starting from an equal sharing of $C$, this mechanism tax every agent that has declared to have at least one preceding task to pay her own declared duration $r^i(\widehat{e})+1$ times. The agents that have declared not to have any preceding task share equally the total tax paid by the formers, pay a quantity equal to their total declared duration time and receive a positive transfer equal to the total tax paid by her partners in $S(\widehat{e})$ divided by $\#S(\widehat{e})-1$. If just one terminal agent exist, her $\widehat{T}_i$ tax is distributed evenly among the non-terminal agents.

It is easy to check that this is an anonymous mechanism: any permutation of the names of the agents only permute their payoffs and no information about the agents' names is used in the mechanism. It is also an invariant to the project size mechanism, since every agent transfer in every circumstance is proportional to both any project size $C$ and the reported duration of the agents. Moreover, it is always balanced: for any reported $\widehat{e} \in PPE$, adding up the agents' transfers yields: $\sum_{i=1}^{n} w_i(N, \widehat{e}_1, ..., \widehat{e}_n) =$

$$= \sum_{i\notin S(\widehat{e})} \left[ \frac{C}{n} - (r^i(\widehat{e})+1)\widehat{T}_i \right] +$$

$$+ \sum_{i\in S(\widehat{e})} \left[ \frac{C}{n} + \frac{\sum_{j\notin S(\widehat{e})}(r^j(\widehat{e})+1)\widehat{T}_j}{\#S(\widehat{e})} - \widehat{T}_i + \frac{\sum_{j\in S(\widehat{e})/\{i\}}\widehat{T}_j}{\#S(\widehat{e})-1} \right] = C \text{ if there}$$

are at least two terminal agents and

$$\sum_{i=1}^{n} w_i(N, \widehat{e}_1, ..., \widehat{e}_n) = \sum_{i\notin S(\widehat{e})} \left[ \frac{C}{n} - (r^i(\widehat{e})+1)\widehat{T}_i + \frac{\widehat{T}_k}{n-1} \right] +$$

$$+ \left[ \frac{C}{n} + \sum_{j\notin S(\widehat{e})}(r^j(\widehat{e})+1)\widehat{T}_j - \widehat{T}_k \right] = C \text{ if just one agent -}k\text{- is terminal.}$$

To prove that it implements the efficient critical path $PPF$, we will compare the payoff each agent obtains by both reporting the truth and lying for any possible $\widehat{e}_{-i}$.

**Case 1**: $i \notin S(e_i, \widehat{e}_{-i})$. We have to distinguish two cases:

**Case 1.1.**: $\#S(e_i, \widehat{e}_{-i}) > 1$.

In this case, reporting a technologically feasible task duration longer than the true one $T_i$, say $\widehat{T}_i > T_i$ will not affect her classification as $i \notin S(\widehat{e}_i, \widehat{e}_{-i})$, since $\widehat{P}_i = P_i \neq \emptyset$ and $r^i(e_i, \widehat{e}_{-i}) = r^i(\widehat{e})$. Hence, agent $i \in N$, by reporting the truth $T_i$, would obtain the following payoff: $w_i(N, e_i\widehat{e}_{-i}) - \sum_{t=\varphi_i^1(N,e_i\widehat{e}_{-i})-T_i}^{t=\varphi_i^1(N,e_i\widehat{e}_{-i})-1} \beta_i^t = \dfrac{C}{n} - (r^i(\widehat{e}) + 1)\widehat{T}_i - \sum_{t=\varphi_i^1(N,e_i,\widehat{e}_{-i})-T_i}^{t=\varphi_i^1(N,e_i\widehat{e}_{-i})-1} \beta_i^t$

By declaring $\widehat{T}_i = T_i + 1$, agent $i$ is enlarging the critical path in 1 hour and, given its true $\beta_i$, can save $\sum_{t=\varphi_i^1(N,e_i,\widehat{e}_{-i})-T_i}^{t=\varphi_i^1(N,e_i\widehat{e}_{-i})-1} \left(\beta_i^t - \beta_i^{t+1}\right) = \beta_i^{\varphi_i^1(N,e_i\widehat{e}_{-i})-T_i} - \beta_i^{\varphi_i^1(N,e_i\widehat{e}_{-i})} > 0 \quad (1)$ monetary units. The cost of obtaining that benefit is 1 monetary unit, but notice that the benefit can never outweight its cost. Now, notice that by declaring $\widehat{T}_i = T_i + 2$, the benefit will be $\sum_{t=\varphi_i^1(N,e_i,\widehat{e}_{-i})-T_i}^{t=\varphi_i^1(N,e_i\widehat{e}_{-i})-1} \left(\beta_i^t - \beta_i^{t+2}\right) = \sum_{t=\varphi_i^1(N,e_i,\widehat{e}_{-i})-T_i}^{t=\varphi_i^1(N,e_i\widehat{e}_{-i})-1} \left(\beta_i^t - \beta_i^{t+1}\right) +$

$+ \sum_{t=\varphi_i^1(N,e_i,\widehat{e}_{-i})-T_i+1}^{t=\varphi_i^1(N,e_i\widehat{e}_{-i})} \left(\beta_i^t - \beta_i^{t+1}\right) < 2$, which is the cost of declaring the task to be 2 hours longer, and so on, so $\forall \widehat{T}_i > T_i$, agent $i$ -actually, any agent- can never find that lying is more profitable than saying the truth.

By reporting a technologically feasible preceding tasks set $\widehat{P}_i \supset P_i$, agent $i \in N$ cannot change neither her non-terminal category nor the fact that $S(\widehat{e}_i, \widehat{e}_{-i}) > 1$, so it holds that $\forall \widehat{P}_i \supset P_i,\ \ i \notin S(e_i, \widehat{e}_{-i}) \Rightarrow i \notin S(\widehat{e}_i, \widehat{e}_{-i})$ and she can get a benefit of at most a delay of $\sum_{j\in\widehat{P}_i} \widehat{T}_j - \sum_{j\in P_i} \widehat{T}_j \geq 0$. By reporting $\widehat{P}_i = P_i \cup \{k\}$, for any $k \in N \backslash \{i\}$, agent $i \in N$ cannot in any case get a direct benefit as large as $\widehat{T}_i$ -see the argument above-, and this only in the case of being part of the critical path after the lie -and maybe before the lie-, in which case $r^i(\widehat{e}) \geq r^i(e_i, \widehat{e}_{-i}) + 1$. Notice that both arguments are valid for possible lies that mix both declaring longer duration time and a larger preceding tasks set.

Finally, revealed $\widehat{\beta}_i$'s do not enter into the definition of the mechanism, so there is no point in lying about it. Notice, however, that no mechanism implementing the efficient $CPA$ can make non-trivial use of information about the revealed $\beta$'s.

**Case 1.2.**: $\#S(e_i, \widehat{e}_{-i}) = 1$. In this case, no lie of any form: $\widehat{T}_i > T_i$ or $\widehat{P}_i \supset P_i$ can change the facts of $i \notin S(\widehat{e}_i, \widehat{e}_{-i})$ and $\#S(\widehat{e}_i, \widehat{e}_{-i}) = 1$, so a transfer of the form: $w_i(N, \widehat{e}_1, ..., \widehat{e}_n) = \dfrac{C}{n} - (r^i(\widehat{e}) + 1)\widehat{T}_i + \dfrac{\sum_{j\in S(\widehat{e})} \widehat{T}_j}{n-1}$ is unavoidable. The third term is independent of the reported lie, so an identical reasoning to that of Case 1.1. applies to this case and there are no incentives to lie.

**Case 2**: $i \in S(e_i, \widehat{e}_{-i})$.

In this case, if agent $i$ is terminal for that technology, any lie consisting in declaring a larger duration $\widehat{T}_i > T_i$ cannot neither change her terminal status nor alter the set $S(e_i, \widehat{e}_{-i}) = S(\widehat{e}_i, \widehat{e}_{-i})$, but can at most delay the project at most $\tau = \widehat{T}_i - T_i$ hours. The benefit for one hour delay is given by (1) and the same reasoning ensures that no such lie can ever yield a direct benefit of $\min\{\tau, T_i\}$ monetary units. In both cases of $\#S(e_i, \widehat{e}_{-i}) > 1$ or $\{i\} = S(e_i, \widehat{e}_{-i})$, since both $r^j(\widehat{e})$ and $\widehat{T}_j$ for all $j \in N \backslash \{i\}$ cannot change with any technologically admissible lie, the cost of the lie is exactly $\widehat{T}_i > T_i$ monetary units -see the payoff function-, which is always bigger than the cost.

If agent $i \in N$ declares to have some preceding tasks, she will always change her own terminal status to non-terminal, so $\forall \widehat{P}_i \supset P_i = \emptyset, \;\; i \in S(e_i, \widehat{e}_{-i}) \Rightarrow i \notin S(\widehat{e}_i, \widehat{e}_{-i})$. There are again two possibilities here:

**Case 2.1.**: $\{i\} = S(e_i, \widehat{e}_{-i})$.

If $S(e_i, \widehat{e}_{-i})$ is a singleton -i.e., $i$ is the only terminal agent for $e_i$ and $\widehat{e}_{-i}$-, there is no technologically feasible lie available to agent $i$ by declaring a non-empty $\widehat{P}_i$, since for any $k \in N$ such that $k \in \widehat{P}_i$, there will necessarily be a sequence of tasks $j, l, z \in N$ such that $j \in \widehat{P}_k, l \in \widehat{P}_j, ... \; z \in \widehat{P}_i$ and we get a cycle, so there is no possibility in this case of getting the non-terminal agents payoff.

**Case 2.2.**: $\#S(e_i, \widehat{e}_{-i}) > 1$. In this case, agent $i$ may get a non-terminal agent status by declaring $\widehat{P}_i \neq \emptyset$ and $\widehat{P}_i \subseteq S(e_i, \widehat{e}_{-i}) \backslash \{i\}$. There are two possibilities now:

**Case 2.2.1.**: $\#S(\widehat{e}_i, \widehat{e}_{-i}) > 1$. In this case, observe that for every possible $\widehat{e} \in PPE$, the total payoff of any terminal agent is always bigger than that of any non-terminal agent, i.e., the following holds:

$$\frac{C}{n} + \frac{\sum_{j \notin S(\widehat{e})} (r^j(\widehat{e}) + 1)\widehat{T}_j}{\#S(\widehat{e})} - \widehat{T}_i + \frac{\sum_{j \in S(\widehat{e})/\{i\}} \widehat{T}_j}{\#S(\widehat{e}) - 1} > \frac{C}{n} - (r^i(\widehat{e}) + 1)\widehat{T}_i. \; (2).$$

Notice that, although the first term is cancelled in both sides, since $r^i(\widehat{e}) \geq 1$ for any non-terminal agent, the second term is always at least $\widehat{T}_i$ monetary units bigger on the right than on the left -in absolute terms-, while the third term on the left is always positive, so by lying declaring a non-empty $\widehat{P}_i$, the cost in terms of the transfer is always bigger than $r^i(\widehat{e})\widehat{T}_i$. On the other hand, the direct benefit of getting a delay of at most $\sum_{j \in \widehat{P}_i} \widehat{T}_j$ -the maximum possible for some $\widehat{e}_{-i}$- is again always bounded by $\min\left\{\sum_{j\in\widehat{P}_i}\widehat{T}_j, T_i\right\}$ monetary units, and the loss in terms of the transfer will be -as was argued above- bigger than $r^i(\widehat{e})\widehat{T}_i$, so since $r^i(\widehat{e})\widehat{T}_i > \min\left\{\sum_{j\in\widehat{P}_i}\widehat{T}_j, T_i\right\}$ for $\widehat{T}_i > T_i$, true terminal agent $i$ has no incentive to declare to be non-terminal.

**Case 2.2.2.**: $\#S(\widehat{e}_i, \widehat{e}_{-i}) = 1$. In this case, agent $i$ lies in the following way: there are other terminal agent initially and by reporting this agent to precede herself makes this agent the only terminal agent after the lie. Henceforth, agent $i$ changes status from terminal to non-terminal and gets an additional bonus by getting a share of the terminal agent tax -see *Figure 1* below-. Agent $i$ will not have an incentive to lie if, for every $\widehat{e} \in PPE$ and $e_i$, the transfer obtained by reporting her true technology -left hand side of inequality (3) below- exceeds the transfer obtained by lying plus the maximum possible direct gain from lying, i.e., $T_i$ -in brackets below- as we have seen above, so if the new terminal agent is $k \in N$, $i \in N$ will not lie if:

$$\frac{C}{n} + \frac{\sum_{j\notin S(\widehat{e})}(r^j(\widehat{e})+1)\widehat{T}_j}{2} - \widehat{T}_i + \widehat{T}_k \geq \frac{C}{n} - (r^i(\widehat{e})+1)\widehat{T}_i + \frac{\widehat{T}_k}{n-1} + [T_i]\,.$$

(3).

If $n = 2$, expression (3) is true when $-\widehat{T}_i + \widehat{T}_k \geq -2\widehat{T}_i + \widehat{T}_k + T_i \Rightarrow \widehat{T}_i \geq T_i$, which is always true by assumption. If $n > 2$, for any $\widehat{e} \in PPE$, the left hand side of (3) becomes bigger as the second term is positive and the right hand side becomes smaller as the second term can only amount to either $-2\widehat{T}_i$ or less -until reaching a minimum of $-(n-1)\widehat{T}_i$-. Moreover, the third term on the right hand side becomes always smaller as the number of agents increase. Obviously, expression (3) holds for every possible economy and lie compatible with the case.

Since we have checked that in every possible economy no lie can be ever profitable, the mechanism is strategy-proof and the proof is complete.

■

Given the above possibility result, the next obvious step is to refine the set of desirable mechanisms by imposing additional properties and test the robustness of the above result to the introduction of other desirable properties in this context. We opt for individual rationality since it is likely to be important in real-life situations. Our next result proves the impossibility of designing anonymous, balanced and individually rational mechanism implementing the efficient critical path $PPF$.

**Theorem 86** *There does not exist anonymous, balanced and individually*

Figure 1:

*rational mechanisms implementing the efficient critical path PPF for any C and $\overline{U}$.*

**Proof.** Let us fix any reservation utility threshold $\overline{U} \in Z$ and any project size $C > 0$. Let us call $g(x) \in Z$ $\forall x \in E$ to the function that assigns the minimum integer between the two closest integers to any real number -the smallest integer between the two closest-. We assume that a mechanism is balanced, anonymous and strategy-proof and will prove that no such a mechanism can ever be individually rational as well. Now, we consider two cases:

**Case 1**: $\frac{C}{2} < \overline{U} + 1$. In this case, consider the following admissible economy $(N, e)$: $N = \{1, 2\}$, $P_i = \emptyset$, and $T_i = 1$ $\forall i = 1, 2$, regardless of the agents' $\beta$'s. By anonymity and balance, $w_i(N, e) = \frac{C}{2}$, so the total payoff each agent receive is $\frac{C}{2} - 1$, which is strictly smaller than $\overline{U}$ by assumption, so individual rationality is violated in this case.

**Case 2**: $\frac{C}{2} > \overline{U} + 1$. We have to prove that within that range of the parameters $C$ and $\overline{U}$, we can find an economy such that individual rationality is violated, provided that we work within anonymous, balanced and strategy-proof mechanisms. Let us consider the following economy $(N, e)$ : $N =$

$\{1,2\}$ , $P_i = \emptyset$, $T_i = g(C - 2\overline{U})$ and $\beta_i$ sufficiently close to 1 $\forall i = 1, 2$. Notice that since $\frac{C}{2} > \overline{U} + 1$ within the assumption, $T_i \geq 2$ and is an admissible integer duration, so $e \in PPE$. Now, let us consider any economy $(N, e')$ such that the only change with respect to $e \in PPE$ is $T_1'$ being very large -tending to infinity-. This economy is also feasible. Observe now that agent 1 in the true economy $e$ will lie and declare $T_1'$ if the total payoff obtained by agent 1 after the transfer is made in $e' \in PPE$ is greater than her total payoff in economy $e$, and since we assume the mechanism to be strategy-proof, we have to impose the following condition to hold: $w_i(N, e_i', e_{-i}) - \sum_{t=\varphi_i^1(N,e_i',e_{-i})-T_i}^{t=\varphi_i^1(N,e_i',e_{-i})-1} \beta_i^t \leq w_i(e) - \sum_{t=\varphi_i^1(e)-T_i}^{t=\varphi_i^1(e)-1} \beta_i^t$. Note that since $\beta_1$ is sufficiently close to 1 and $T_1 < \infty$ but still $T_1'$ is infinitely longer than $T_1$ by construction, the expression above can be written as follows for some appropriate selection of both $T_1'$ and $\beta_1$ :

$$w_1(N, T_1', \emptyset, e_2) - \varepsilon_1 \leq \frac{C}{2} - T_1 - \varepsilon_2$$

for some $\varepsilon_1, \varepsilon_2 > 0$ but as close to 0 as desired, so there exist admissible economies such that strategy-proofness, anonymity and balance require some $w_1(N, T_1', \emptyset, e_2) - \varepsilon_1$ to be smaller than $\frac{C}{2} - T_1$. Substituting $T_1$, we obtain: $w_1(N, T_1', \emptyset, e_2) - \varepsilon_1 \leq \frac{C}{2} - g(C - 2\overline{U}) < \frac{C}{2} - (C - 2\overline{U})$. Finally, notice that the right hand side of the last inequality can be written as $\left[\overline{U} - \frac{C}{2}\right] + \overline{U}$, which is always smaller than $\overline{U} - 1$ under our assumptions, so the following holds: $w_1(N, T_1', \emptyset, e_2) - \varepsilon_1 < \overline{U}$. But notice that $w_1(N, T_1', \emptyset, e_2)$ still applies if the true economy was $e'$, in which case the direct cost of undertake the true task $T_1'$ on the left hand side of the last inequality will be much larger than the negligible $\varepsilon_1$, so for that true economy the net agent 1's payoff will be much smaller than $\overline{U}$, so no mechanism can be individually rational for any $\overline{U}$ and any $C > 0$. ■

## 5.4 Concluding Remarks

In this chapter we have explored the possibility of designing strong incentive compatible mechanisms in a particular production setup: the well-known network production problem, and to the solutions defined by methods like the CPM and the PERT when transfers to the agents are possible. Although the problem is quite similar to other environments like the public goods problem, the opportunities for exploiting the asymmetry on the distribution of the private information is very different in this context, and it is not surprising that we get different results of those of Groves and Clarke -see, for example, Groves [4] and Groves [5]- in the public goods problem. Assuming that the agents' payoffs are quasi-linear on the part of the transfers, we find simple strategy-proof, anonymous and invariant to the project size mechanisms implementing the PERT that are balanced as well, so complete efficiency is achieved provided that we include the planner in the definition of Pareto-optimality. Furthermore, if we add other plausible mechanism feasibility property like individually rational payoffs, we obtain an impossibility result. The possible ways-out in this case include imposing constraints on the domain of possible economies or relaxing the equilibrium concept used, although perhaps the most promising approach to escape from the impossibility could be imposing reasonable bounds on the technologies allowed to be considered by the planner, like some maximum time for any task to be completed. The nature of the proofs -and this also includes the Groves-Clarke mechanism in the public goods provision problem- points to this lack of bounds as the key factor.

# References


[1] Bazawa, M.S. and Jarvis, J.J.: "Linear Programming and Network Flows", 1990, John Wiley, New York.

[2] Derigs, V.: "Programming in Networks and Graphs", 1988, LNEMS, 300, Springer, Berlin.

[3] Deo, N. and Pang, C.: "Shortest Path Algorithms: Taxonomy and Annotation", 1984, Networks, 14.

[4] Groves, T.:"Incentives in Teams," Econometrica, 41 (1973), 617-631.

[5] ———:"Efficient Collective Choice when Compensation is Possible," Review of Economic Studies, (1979), 227-241.