FOUR ESSAYS ON IMPLEMENTATION THEORY

Ville Korpela

Sarja/Series A-5:2011

Turun kauppakorkeakoulu
Turku School of Economics

ISBN 978-952-249-166-4 (nid.) 978-952-249-167-1 (PDF)
ISSN 0357-4652 (nid.) 1459-4870 (PDF)
UDK 330.105
330.83
330.36
519.86
519.83

Uniprint, Turku 2011

Acknowledgements

This thesis was written while I was affiliated as a junior researcher with the Public Choice Research Centre (PCRC), one of the Centres of Excellence in research (CoEs) funded by the Academy of Finland (2008-2014). An extensive work like this involve numerous people that deserve a special thanks.

To begin with, I want to thank my thesis supervisor Hannu Vartiainen. The guidance of Hannu was so excellent that any attempt to dress my gratitude into words would hardly seem like enough. The truth is that, without his input, patience and deep knowledge of the subject, this thesis would not exist. I also want to thank Juuso Välimäki and Klaus Kultti, the two pre-examiners of this thesis, for all the valuable comments they have given me. Furthermore, I want to express my sincerest gratitude to Tomas Sjöström for accepting the invitation to act as an opponent for this thesis.

It is impossible to thank everyone that has given me an advise, or another, during these years. Therefore, I simply want to thank the entire staff at Turku school of Economics and at PCRC for providing a good working atmosphere. In particular, I want to thank Hannu Salonen for helping me with some concepts of game theory that I struggled with, and also for reading through some of the papers; Matti Virén who helped me a lot at the time I was taking my first steps as a graduate student; Hannu Nurmi and Juha Tervala for giving me comments on some of the papers; and both Jouni Jokinen and Henna Mäki for helping me with practical details too numerous to remember.

Anyone that has written a thesis knows the importance of peer support. For this reason, among others things, I want to thank Tomi Kortela, Oskari Vähämaa, Pellervo Hämäläinen, Olli Lappalainen and Hannu Autto.

Discussions with you, on and off the subject, are always fascinating. In particular, I would like to thank Tomi Kortela for helping me with the LaTeX and some other technicalities.

Last, but by no means least, I want to thank my mate Elina and my parents Lasse and Raija for all the love and care. I am deeply sorry about the neglect you had to experience during this project. Finally, I want to thank Turun kauppakorkeakouluseura for the financial support.

Turku, August 2011

Ville Korpela

Contents

Part I

A Brief History of Mechanism Design and Implementation Theory

THIS INTRODUCTORY PART is a quick look at the research papers in Part II of this thesis from a broad perspective. This is not a summary of the papers. The main purpose is to place the studies in their proper historical context. However, a short abstract is available at the beginning of each paper. Although, there is repetition in this thesis, it is unavoidable, since the studies are intended to be individual research papers.

Historically, one can separate two lines of research that have both led to the emergence of implementation theory and mechanism design. We present these in turn in chapters 1 and 2. Then, in chapter 3, we concentrate on the current state of the theory and elaborate on what is considered important right now.

To avoid unnecessary repetition, references at the end of each research paper are indicated by an additional number identifying the paper (1, 2, 3 or 4). For example, if we want to refer to Maskin (1999), which is the main reference of the second research paper, we write Maskin (1999-2).

1 The Birth of Social Choice Theory

The influential theorem of Kenneth Arrow (1951), on the problem of aggregating individual preferences into a social ranking, is generally considered to mark the beginning of mathematical social choice theory. This important theorem has generated a whole research agenda and it is well beyond the scope of this introduction to review even the most essential part of the literature (see Arrow, 1951; Fishburn, 1973; Plott, 1976-3; or Sen, 1986, for example). Therefore, to get started, we simply present a very basic version of this theorem.

1.1 Arrow's Impossibility Theorem

Even though we are trying to be as non-technical as possible, some formal definitions will be needed. After all, we are talking about mathematical social choice theory. Let $I = \{1, \ldots, n\}$ be the set of *individuals*, or *agents*, and A the set of *social alternatives.* Every individual $i \in I$ is endowed with a complete and transitive preference relation R_i over A. The set of all possible preference relations over A is denoted by $\mathscr{R}_A$. Occasionally, the designer (an individual or perhaps third party) of a social choice mechanism (e.g., voting rule) does not know the preference relation of individual i exactly. Rather, only the set of possible preference relations $\mathscr{R}_i \subseteq \mathscr{R}_A$ is known. The set $\mathscr{R} = \mathscr{R}_1 \times \cdots \times \mathscr{R}_n$ is called a *preference domain* and $U = \mathscr{R}_A \times \cdots \times \mathscr{R}_A$ is called the *unrestricted preference domain.*[1] A typical element of $\mathscr{R}$ is denoted by $\boldsymbol{R} = (R_1, \ldots, R_n)$, which we call a *preference profile.* As usual, $\boldsymbol{R}_{-i}$ is a preference profile $(R_1, \ldots, R_{i-1}, R_{i+1}, \ldots, R_n)$ that specifies a preference relation for every individual except i.

An "ideal way" of making a *public decision* or a *social choice* would be to find a rule $\Psi : U \to \mathscr{R}_A$, which we may call the *Aggregation Rule.* This rule gives a social ranking $\succeq$ (a complete and transitive binary relation) of the alternatives in A as a function of the preference profile. That is, a rule Ψ such that

$$\Psi(\boldsymbol{R}) \in \mathscr{R}_A \text{ for all } \boldsymbol{R} \in U. \tag{1}$$

After we have found Ψ, social choice is bound to be easy. We simply choose the alternative that is best according to the social ranking $\succeq$.

However, we should require that Ψ satisfies certain natural properties for us to consider it a "reasonable" rule of preference aggregation. After all, not many people would accept the *dictatorial aggregation rule* Ψ_i^D defined as: $\Psi_i^D(\boldsymbol{R}) = \boldsymbol{R}_i$ for all $\boldsymbol{R} \in U$; that is, an aggregation rule that always places the top alternative of individual i as the top alternative of whole society, and which is totally insensitive to the preferences of other individuals. To this end, the following two properties seem harmless enough.

Unanimity (Un): For any $\boldsymbol{R} = (R_1, R_2, \ldots, R_n) \in U$, if $R_i = R$ for all $i \in I$, then we must have $\Psi(\boldsymbol{R}) = R$.

[1] In the most general case, we could assume $\mathscr{R} \subseteq \mathscr{R}_A \times \cdots \times \mathscr{R}_A$. This means that there is a logical connection between the preferences of different individuals.

Independence of Irrelevant Alternatives (IIA): For any $a, b \in A$ and any $\boldsymbol{R} \in U$, if $a\boldsymbol{R}_i b$ for all $i \in I$, then it must also be that $a\Psi(\boldsymbol{R})b$.

The first property is certainly indisputable; whereas the second property is heavily criticized in the literature. *Borda count*, for example, does not satisfy IIA.[2] Still, even IIA seems like a reasonable assumption. Nevertheless, together these properties force us into an unsettling conclusion.

Arrow's Impossibility Theorem. Let $|A| \geq 3$ and assume that the aggregation rule $\Psi : U \to \mathscr{R}_A$ satisfies both Un and IIA. Then, it must be that $\Psi = \Psi_i^D$ for some $i \in I$. ■

Instead of the idealized way of making the social choice put forward in equation (1), which produces negative results, we could take a much more "pragmatic view" on the problem. Even though there is no good way of aggregating individual preferences into a social ranking, at the end of the day, we have to make a choice. But we do not need a social ranking $\succeq$ to make a choice; we need only the socially best alternative. In fact, and to be even more pragmatic, we need only a socially acceptable alternative. This suggests a new formulation of the problem put forth in equation (1). Let $f : U \to A$ be a function that represents the goal of society, and let us call it a *Social Choice Rule* (SCR). This function (or more generally, a correspondence) connects a socially acceptable alternative to every admissible preference profile. In this formulation, we implicitly assume that the designer of the SCR does not know *a priori* what is the truly prevailing preference profile. This is why the domain of the SCR f is set to be U. Unfortunately, this pragmatic formulation will also force us to a negative conclusion, albeit of a different sort.

[2] We will be speaking about voting rules that we do not define. For exact definitions, see Nurmi (1987) or Saari (1994), for example.

1.2 The Gibbard-Satterthwaite Theorem

When we present the social choice problem in the form of equation (1), it is implicitly assumed that the designer can somehow elicit the true preference of all individuals involved. That is, for some reason, all individuals act sincerely or non-strategically. This brings us to the concept of (strategic) manipulation. Consider a presidential election, for example. Assume that f is the majority voting rule. This means that majority voting resolves social acceptability. There are three candidates, $A = \{a, b, c\}$, and with a little abuse of notation, 45% of the individuals have the preference $aPbPc$, 35% have the preference $bPaPc$ and the remaining 20% have the preference $cPbPa$. Now assume that there is a poll and this information becomes common knowledge.[3] Consequently, the 20% of individuals who prefer c will (presumably) vote for b instead. Why is this so? Well, the 20% of individuals know that whatever will happen, none of the remaining 80% will vote for c, and hence, this candidate has no chance to win. Naturally, from the two remaining candidates, these individuals prefer to vote for the better one. But the 35% of individuals who prefer candidate b know this. Therefore, candidate b will be selected at the end. Even though this is not what the majority voting rule would dictate.

We are not trying to make a normative judgement on whether b should be selected in this case or not. The main point of this example, for now, is to show that the phenomenon called manipulation does not necessarily vanish in a referendum context, when there are only few alternatives to choose from (the number of individuals is immaterial in the above example). The most natural way to formally define what is meant by a non-manipulable SCR is strategy-proofness (see Barberà, 2001-4, for example). In this general definition, the domain of an SCR does not have to be the unrestricted domain U.

Strategy-Proofness.[4] An SCR $f : \mathscr{R} \to A$ satisfies *strategy-proofness* if,

[3] See Aumann (1976) or Milgrom (1981) for a definition of common knowledge event.

[4] Here, $\boldsymbol{P}_i$ is the strict part of $\boldsymbol{R}_i$.

for all $\boldsymbol{R} \in \mathscr{R}$ and all $i \in I$, there is no $R'_i \in \mathscr{R}_i$ such that

$$f(R'_i, \boldsymbol{R}_{-i})\boldsymbol{P}_i f(\boldsymbol{R}).$$

Strategy-proofness is a strong incentive compatibility condition. It says that no agent can gain by lying regardless of whether everyone else is speaking the truth or not. But the idea behind strategy-proofness is much more subtle than this. Let us return to the example of a presidential election. How many people believe that since voting is the cornerstone of democratic society, we must be sincere in voting to guarantee that the right outcome prevails? The answer is practically no one. Quite the contrary. We tend to think that we are free to express our will by voting in any way desired. This means that a voting rule that is not strategy-proof does not really produce the outcome it is supposed to.

Unfortunately, the requirement of strategy-proofness will also lead us to negative results (Gibbard, 1973-1; Satterthwaite, 1975-1). We need the following concept. An SCR $f : \mathscr{R} \to A$ is called *dictatorial* if there exists an individual $i \in I$, such that $f(\boldsymbol{R}) \in \{a \in A \mid a\boldsymbol{R}_i b \text{ for all } b \in A\}$ for all $\boldsymbol{R} \in \mathscr{R}$. If this is the case, we denote $f = f_i^D$.

The Gibbard-Satterthwaite Theorem.[5] Assume that $|A| \geq 3$ and let $f : U \to A$ be strategy-proof and onto (surjective). Then, it must be that $f = f_i^D$ for some $i \in I$. ∎

What can we do? It seems that we constantly end up with these dictatorial rules that are badly at odds with the idea of democracy. Three ways to avoid the conclusion of the Gibbard-Satterthwaite theorem come up immediately. First, the choice could be made between only two alternatives, so that $A = \{a, b\}$. Second, we could assume that the designer has information, or that she can at least acquire some. That is, the domain of f could be restricted. Third, and this finally takes us to mechanism design and implementation theory, we could assume that the SCR f is not itself the mechanism used to implement it (an apology is in order, we are using a lot of words not yet

[5] *Plurality voting* rule works as evidence that the assumption $|A| \geq 3$ is necessary. For more on the Gibbard-Satterthwaite theorem, see Danilov and Sotskov (2002), for example

formally defined).

Let us once again return to the example of a presidential election. There is a very simple explanation why we should expect to see a phenomenon called manipulation in this voting situation. Namely, voters do not consider majority voting rule as something normative, representing a core idea of democracy. Rather, they consider it a social choice mechanism, and their own vote a strategy that they can freely choose.

In the parlance of game theory, a *mechanism* or a *game form* is a tuple $G = (\Sigma, g)$, where $\Sigma = S_1 \times \cdots \times S_n$ is the *strategy space* and $g : \Sigma \to A$ is the *outcome function.* For a fixed preference profile $\boldsymbol{R} \in \mathscr{R}$, this mechanism defines a game $\Gamma(\boldsymbol{R}) = (G; \boldsymbol{R})$ in normal form. Let S be a solution concept.[6] All solutions in the game $\Gamma(\boldsymbol{R})$ with respect to S are denoted by $S\big[\Gamma(\boldsymbol{R})\big] \subseteq \Sigma$. We say that *mechanism G implements the SCR $f : \mathscr{R} \to A$ in S* if

$$g\left(S\big[\Gamma(\boldsymbol{R})\big]\right) = f(\boldsymbol{R}) \text{ for all } \boldsymbol{R} \in \mathscr{R}. \tag{2}$$

That is, the outcomes of the mechanism g at the solutions defined by S must exactly coincide with the SCR f. If there exists a mechanism that implements f in S, then we say that *f is implementable in S.*

We will cite equation (2) as the *implementation equation* or the *implementation problem.* This formula is an equation in the sense that mechanism g is an unknown variable. It was Hurwicz (1960,1972) who first formulated the idea that institutions should be the design variable, but this was done only in the context of resource allocation rules. Notice that, in contrast to the standard problems of game theory, where we are trying to predict an outcome of a given game, equation (2) is a kind of "reverse engineering" problem. Now the equilibria are given and we are trying to find the game. Using these formal definitions we can explain the implicit restriction behind the Gibbard-Satterthwaite theorem. Simply, it assumes that $\Sigma = U$ and

[6]This solution concept is not necessarily an equilibrium in any game theoretic sense. This is why Jackson (2001) suggested that we use the term "solution concept" instead of the term "equilibrium concept" .

$g = f$, which is the same as saying that the SCR f is itself the mechanism used to implement it. Unfortunately, this is all about to backfire − again.

1.3 The Revelation Principle

A substantial part of the solution concepts that are analyzed in game theory satisfy the so called *revelation principle.* Dominant strategy equilibrium (Gibbard, 1973-1; Satterthwaite, 1975-1), Bayesian-Nash equilibrium (Myerson, 1979-4; Harris and Townsend, 1981-4) and maximin strategies (Thomson, 1979), for example, all satisfy this principle.[7]

The mechanism $G = (\mathscr{R}, f)$, which we encountered at the end of the previous section, is called a *direct revelation mechanism.*[8] We say that SCR f is *truthfully implementable in* S if truth-telling is a solution of the direct revelation mechanism (with respect to S). Note that, according to this definition, other solutions may exist, besides the truth-telling one. The main point here is that if S is chosen as one of the soultion concepts mentioned earlier, then the following theorem holds true.[9]

Revelation Principle. If SCR $f : \mathscr{R} \to A$ is implementable in S, then it is also truthfully implementable in S. ■

So, why is this result negative? Some could claim that the result is in fact positive, since now we can restrict attention in the set of direct revelation mechanism only. At least, if we are willing to accept truthful implementation as satisfactory. To understand the negative connotation of this result, consider implementation in dominant strategy equilibrium. This is, in many ways, the most natural solution concept that we can use in an implementation problem. Assume that mechanism g implements the SCR f in dominant strategies. Then, according to the revelation principle, f must also be truthfully implementable in dominant strategies. But if truth-telling is a dominant strategy equilibrium of the direct revelation mechanism, then f

[7]One notable exception is Nash equilibrium. We discuss this solution concept later on.
[8]Earlier we had $\mathscr{R} = U$, but this does not have to hold in general.
[9]See Dasgupta et al. (1979-1) for a more detailed discussion of this principle.

must obviously be strategy-proof. Therefore, we are back to the Gibbard-Satterthwaite theorem. This is also true to some extent for all other solution concepts that satisfy the revelation principle – truthful implementation is simply quite restrictive.

All this leaves us with two possible avenues; Either we start to use a solution concept that does not satisfy the revelation principle, so that we can get rid of the dominant position held by direct revelation mechanisms; or we must consider the possibility that the designer of a social choice mechanism can acquire information about the preferences, so that the domain of an SCR f is somehow restricted (excluding the case of A having only two alternatives). In the next section, we concentrate on the latter case. This approach has the added value that we can still use dominant strategy equilibrium as the solution concept.

1.4 VCG Mechanisms and Single-Peaked Preferences

Why should we regard the assumption of unrestricted domain in the Gibbard-Satterthwaite theorem as reasonable? In practice, social choice is mostly done by voting.[10] Assume that the set of social alternatives A does not have any natural structure and that the designer has to choose a voting rule to be used.[11] In this case, it might be reasonable to assume that the designer does not know anything about the possible preferences of the voters. That is, the domain of an SCR (here, the voting rule that resolves social acceptability) should indeed be considered unrestricted.

However, the set of social alternatives A usually does have some kind of logical structure, a structure that is recognizable to the designer by virtue of the problem at hand. Think about the problem of selecting a socially optimal amount of public good, for example. Samuelson (1954) was the first one to recognize that this problem cannot be solved through the market mechanism because of the free rider problem. In this public good setting, as

[10] Think about the decision making of a Parliament.

[11] A could be a metric space, or more generally, a topological space.

in any economic setting, we can use money to create incentives, since it is quite reasonable to assume that everyone prefers more to less. Consequently, we could assume that the domain of an SCR that assigns a socially optimal amount of public good to every admissible preference profile consists of all utility functions that are quasi-linear in money. But then, mechanisms do exist, called VCG -mechanisms (named after Vickrey, 1961; Clarke, 1971; and Groves, 1973), that are strategy-proof and give the socially optimal amount of public good for all admissible profiles of utility functions.[12]

Yet, a similar thing can also happen in a voting context. Assume that the social alternatives in A can be ordered $a_1, a_2, \ldots, a_k$ according to a left-wing – right-wing spectrum. That is, a left-wing member would consider a_1 the best alternative, and a right-wing member would consider a_k the best alternative. Then, it might be reasonable to assume that all voters have single-peaked preferences over A. Basically, this means that all voters have a best alternative $a_j \in A$, called the "peak," and the closer another alternative $a_l \in A$ is to this one (in the sense of $|j - l|$ being smaller), the better it is considered to be. Let the utility functions $u_1, \ldots, u_n$, one for each voter, represent single-peaked preferences over A. Denote the peak of u_i by $P(u_i)$. Moulin (1980-1, 1983-1) has shown that the *Generalized Median Voter Rule* (GMVR)

$$f(u_1, \ldots, u_n) = \text{median of} \{P(u_1), \ldots, P(u_n), \alpha_1, \ldots, \alpha_m\},$$

where the numbers $\alpha_1, \ldots, \alpha_m$, n and m must satisfy certain restrictions that we do not discuss here, is strategy-proof over any domain that contains only single-peaked preferences.

Now we know that strategy-proof SCRs do exist on restricted domains that can be considered realistic. In fact, the search for new domains that allow strategy-proof SCRs is very much alive today. The most obvious advantage of using strategy-proof SCRs is that the search for incentive compatible mechanisms can be restricted in the set of direct revelation mechanims.

[12]These mechanisms have certain well-known problems with budget balance that we do not discuss here, see Roberts (1979), for example.

Unfortunately, things are a bit more complicated than this. Namely, direct revelation mechanisms can be informationally very inefficient.

An example will illustrate this. Assume that $\mathscr{R}_i$, insted of being a set of preference relations, is now the set of all utility functions that represent a single-peaked preference over the set $A = \{a, b, c, d\}$. The designer has chosen GMVR as the SCR. Direct revelation mechanism would require all individuals to send a vector of real numbers, such as $[u(a), u(b), u(c), u(d)] = (1.5, 3, 2, 0.5)$, to the designer.[13] There are, of course, an infinite amount of these. But GMVR does not use all the information that is in this vector of real numbers. It will only use the fact that b is the peak. One can easily see that an indirect mechanism that would ask every individual to submit only the peak, or to vote for the most preferred alternative, would produce exactly the same outcome. Moreover, the informational requirements of this indirect mechanism would be considerably smaller.

This takes us to the topic of the **first research paper**. In this paper we try to find out the minimal number of strategies that are needed to implement an SCR in dominant strategies or securely.[14] In both of these cases, we know, by the revelation principle, that the direct revelation mechanism implements the SCR truthfully. But as the example in the previous paragraph shows, we can dig deeper. We can ask whether an indirect mechanism could do any better. The main result is that, quite generally, the number of strategies that are needed in a secure- or dominant strategy implementation of an SCR is exactly the same as the number of messages that are needed in a decentralized mechanism realizing the SCR (these concepts are defined at the beginning of the first research paper).[15]

At this point, we can also explain the idea of the **fourth research paper**,

[13]Note that this represents a single-peaked preference.

[14]*Secure implementation* is a term coined by Saijo et al. (2007-1). This is a form of *double implementation*, in which the solution concepts are Nash and dominant strategy equilibrium.

[15]Williams (2008) is an in depth treatment of decentralized mechanisms. We have interest on this topic merely as a by product.

even though to fully appreciate it we must return to it later on. There has been a growing trend in implementation theory to incorporate more behavioral assumptions into the theory (see Matsushima, 2008a-4 and 2008b-4, for example). At the same time, there is nothing in the implementation problem (2) that would force us to make the prevailing assumption of complete and transitive preferences. Assuming, of course, that we can define a suitable solution concept S to fit the more general framework. After this has been done, we must ask whether we can still use the revelation principle. That is, what are the behavioral boundaries of the revelation principle. It turns out that we do not need to assume full rationality, only a well-known condition called Property α (Sen, 1977-3).

2 The Debate Between the Relative Merits of Socialist Economies versus Market Based Economies

The other line of research that led to mechanism design, and then to implementation theory, has its roots in the debate between the relative merits of socialist economies versus market based economies that took place in 1930s and 1940s. In fact, this debate is why the formal theory put forth in equation (2) was originally created.

2.1 The Hayek-Mises-Lange-Lerner Debate

In 1930s and 1940s there was a lengthy discussion about the feasibility of socialist economies (see Hayek, 1935; Lange, 1937; Lerner, 1936; and von Mises, 1920). This was later to become known as the Hayek-Mises-Lange-Lerner debate (Moore, 1992). On a large scale, the question was already about mechanism design: Which is the better resource allocation mechanism – market economy or economy based on planning. It was Hayek (1945) who first expressed the most elegant argument for market economy ever given: Since "*the data from which the economic calculus starts are never for the whole society given to a single mind*", the problem to be solved is "*how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know*" [reproduced from Maskin and Sjöström, 2002]. That is, an economy based on planning assumes that one man could possess all the relevant information, while market economy does the resource allocation in a decentralized way. However, there was something that everyone participating in this debate had overlooked.

2.2 Hurwicz and the Resource Allocation Problem

In two notable papers, Hurwicz (1960,1972) took a much more general view on the resource allocation problem. The starting point was an observation about the current state of the debate: Who is to say that neither one, market economy or economy based on planning, is the best resource allocation mechanism. After all, what we see in practice is something in between. This led Hurwicz to formally define the idea of a mechanism, and makes him the father of mechanism design and the forefather of implementation theory.

This general formulation left economists pondering what is the truly distinctive feature of market economy. That is, what distinguish market mechanism from all other resource allocation mechanisms. The fact that it is decentralized is obvious. On the other hand, there are certainly other decentralized mechanisms (maybe even an infinite amount of them). The first big breakthrough of mechanism design, as a field of its own, was to show that market mechanism is informationally the most efficient way to obtain Walrasian allocation (see Osana, 1978-1, for example).

2.3 The Genesis of Implementation Theory

It was Erik Maskin (1977-3, 1999-2) who first formulated the implementation problem (2) for an arbitrary SCR f, not just for resource allocation rules, using a solution concept S that does not satisfy the revelation principle. The solution concept that he used is the single most important idea in modern game theory, that of Nash equilibrium (see Kuhn and Nasar, 2002, for example).

To reiterate, the research tradition that followed Arrow (1951) never ended up considering any solution concepts that do not satisfy the revelation principle. In fact, the most common notion of incentive compatibility was simply that of strategy-proofness. This is the reason that SCR was itself considered the mechanism that is used to implement it.[16] On the other hand, the research tradition that began from the debate between the relative merits of socialist economies versus market based economies never went beyond resource allocation rules. Therefore, the paper of Maskin (1977-3) must be

[16] The "sophisticated voting" of Farquharson (1961) is somewhat of an exception, but there are certainly others too.

considered the true genesis of implementation theory. For this, among other things, he was awarded the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel together with Leonid Hurwicz and Roger Myerson in 2007.[17]

The result of Maskin (1977-3) states that *monotonicity* is a necessary condition for an SCR to be implementable in Nash equilibrium, and together with a condition known as *no-veto-power*, it is also a sufficient condition for Nash implementation.[18] Since this paper there has been an ongoing search for one condition that would be a full characterization (a necessary and sufficient condition). Danilov (1992-2) and Yamato (1992-2) solve this problem for a restricted class of preference domains. Bochet (2007) solves the problem when the designer can use lotteries over the social alternatives A. The result of Bochet (2007) is quite surprising, monotonicity alone is a full characterization. In the **second research paper** we give yet another solution to this problem. We show that a generalization of monotonicity exists that is a full characterization without any domain restrictions or otherwise special structure (*i.e.*, whether lotteries are possible or not).

We also work closely in this "Maskinian tradition" in the **third research paper**. As we already mentioned, there is nothing in the implementation problem (2) that would force us to assume that choice behavior is generated by a complete and transitive preferences relation. It was Hurwicz (1986-3) who first observed this fact. The most natural way to formulate the implementation problem (2) without the assumption of rationality is to start from choice functions. This requires that the solution concept S must also fit this more general choice function setting. In the **third research paper** we claim that the solution concept used by Hurwicz (1986-3), called the *Generalized Nash Equilibrium*, is not entirely consistent with the assumption

[17]We must stress that the use of Nash equilibrium does not free us totally from the problems that plague truthful implementation (or dominant strategy implementation). This is due to the fact that informational assumptions of the two solution concepts, Nash equilibrium and dominant strategy equilibrium, are different. However, the results in Nash implementation do show that sometimes we can accomplish more.

[18]See the second research paper for exact definitions.

that behavior is describable by a choice function. If we use this solution concept, then we have to assume that choice behavior is *normal* (see Sen, 1977-3). In the second part of the paper we give a new solution concept that better fits this framework. Using this, we show that the old results in Nash implementation theory still hold if we assume only a condition known as Property α (rationality of behavior also requires another condition, called Property β). We also show, through an example, that if Property α does not hold, then the old characterization results do not necessarily hold either.

We can now finally explain also the deeper idea behind the **fourth research paper**. Most characterization results in implementation theory are obtained in the following way (in a technical sense): First, verify that truth-telling is an equilibrium of the direct revelation mechanism. There might be, and usually also are, other equilibria. Second, augment the direct revelation mechanism with additional strategies to obtain uniqueness, or at least, to remove all non-desirable equilibria. Since the behavioral assumption that we need for the revelation principle to still work is Property α, we can expect that it is important in every type of implementation and will come up repeatedly. This property is a kind of wedge between being able to use only augmented direct revelation mechanisms and having to enter the world of truly indirect mechanisms.[19]

[19] Mookherjee and Reichelstein (1990) give a formal definition of an augmented direct revelation mechanism.

3 Modern Era: It Boils Down to "Social Engineering"

After the groundbreaking result of Maskin (1977-3), the field expanded rapidly. Therefore, it is not possible to describe all that is on agenda right now. We are only going to present some broad guidelines. There are numerous general treatments that the reader can consult, including Baliga and Sjöström (2007), Chorchón (1996), Chorchón (2007), Jackson (2001), Maskin and Sjöström (2002), Moore (1992) and Serrano (2003).

3.1 Refined Solution Concepts and More Elaborate Mechanisms

Subsequent to solving the implementation problem (2) using Nash equilibrium as the solution concept S, it is only natural to ask what would happen if the solution concept is chosen as some refinement of Nash equilibrium and the mechanism g as an extensive form game (as opposed to a normal form game). An obvious candidate is the *subgame perfect equilibrium* of Selten (1975). This problem was first studied by Moore and Repullo (1988) and Abreu and Sen (1990), and later completely solved by Vartiainen (1999, 2007).[20] The result is striking and, in a sense, transforms the whole field into "social engineering." The condition for subgame perfect implementation turns out to be very weak. This transformation was made even more striking by the results in *undominated Nash implementation* (Palfrey and Srivastava, 1991), *virtual implementation* (Abreu and Sen, 1990) and *implementation in perfect equilibrium* (Sjöström, 1993).[21] Almost nothing but at least one preference reversal between different preference profiles in the domain is needed.

[20] Herrero and Srivastava (1992) solve this problem when only games of perfect information are allowed.

[21] For a criticism of undominated Nash implementation and a resolution, see Jackson (1992) and Jackson, Palfrey and Srivastava (1994), respectively.

However, as the reader may have already realized, things are a bit more complicated: The solution concept is not really a choice variable of the designer. The mere fact that all individuals can select the best alternative from any choice set does not imply that Nash equilibrium, let alone subgame perfect equilibrium, would be played. In other words, the solution concept that we use will unavoidably incorporate a lot of behavioral assumptions. In fact, it obviously requires a higher level of rationality to play a subgame perfect equilibrium than Nash equilibrium. This is why it is important to understand every form of implementation, so that we can choose the right solution concept for the problem at hand. In effect, then, these recent developments in implementation theory have not rendered the old result of Maskin (1977-3) in any way outdated.

3.2 Robust Mechanism Design

So far, we have only talked about implementation under specific informational assumptions. To be more precise, all solution concepts that we have presented assume that everyone knows the whole preference profile (if Nash equilibrium is used) or does not need to know anything about others (if dominant strategy equilibrium is used).[22] We could as well assume that every individual has only a belief about the truly prevailing preference profile. This branch of the implementation literature is known as implementation under incomplete information or *Bayesian implementation* (for the most general result, see Jackson, 1991; for an overview, see Palfrey, 1992). The surprising thing is that the results embody an extensive analogy with the complete information case.[23]

However, the informational assumptions behind Bayesian implementation are much more demanding that those behind other types of implementation.

[22] We are not going to defend these assumption here, even though they are central in this thesis. These assumptions are sometimes realistic and sometimes not (see Moore, 1992, for example).

[23] Notice that Bayesian equilibrium is nothing more than a Nash equilibrium in an incomplete information setting.

This time the burden is just on the designer, since we must assume that the designer knows the beliefs (as agents now maximize expected utility). We do not have a single example in mind that would render this assumption completely realistic. On the other hand, if we require that the mechanism must implement the SCR no matter what the true beliefs are, then an elegant result of Bergemann and Morris (2005-1), on *robust implementation*, shows that we are back with dominant strategy implementation quite tightly.[24] As we have already noted, this is a very restrictive form of implementation. The disappointing thing is that, if we consider implementation under incomplete information as the most realistic framework, then we have made an almost complete circle and returned to the original position, where only direct revelation mechanism where used.

Fortunately, the result of Bergemann and Morris (2005-1) gives a more pessimistic view than what is really justifiable. This is because all beliefs are not logically possible. To push the argument even further, the designer may sometimes be able to affect the beliefs through the mechanism. This should convince the reader that there is a lot of unexplored territory in the literature on Bayesian implementation, and it is anything but obvious where it will lead. One thing is certain, information has and will be an important ingredient in explaining social and economic organization (see Akerloff, 1970; Mirrlees, 1971; and Rothschild and Stigliz, 1976, for example).

3.3 The Prevailing Consequentialist View

To conclude, we just note that the idea behind the implementation equation (2) is very much consequentialist in nature. That is, given any SCR f, or the goal, the implementation problem is solved by any mechanism g satisfying equation (2). But the fact is that people may accept unfair outcomes if they are obtained through a procedurally fair system; or on the other hand, not accept fair outcomes that are obtained through a procedurally unfair system (here, g can be interpreted as the "system"). This observation suggests

[24] See Eliaz (2002) for another notion of robustness.

that there should be some kind of endogeneity between the SCR and the mechanism that is used to implement it. A preliminary step towards solving this type of questions is Jackson and Wilkie (2005).

This is all that we want to say at this point. The central motivation behind the research papers in this thesis should now be clear. One should keep in mind that the “story” we told in this introduction is something that a historian would call a *super rational reconstruction*: We have not made any effort to interpret these past events from a contemporary point of view.

References

Aumann, R. (1976), Agreeing to Disagree. *Annals of Statistics* 4, 1236-39.

Abreu, D. and Sen, A. (1990), Virtual Implementation in Nash Equilibrium. *Econometrica* 59, 997-1021.

——— (1990), Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition. *Journal of Economic Theory* 50, 285-299.

Akerloff, G., A. (1970), The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics* 84, 488-500.

Arrow, K., J. (1951), *Social Choice and Individual Values.* John Whiley, USA.

Baliga, S. and Sjöström, T. (2007), Mechanism Design: Recent Developments. *The New Palgrave Dictionary of Economics*, Second Edition, edited by Larry Blume and Steven Durlauf.

Bochet, O. (2007), Nash Implementation with Lottery Mechanisms. *Social Choice and Welfare* 28, 111-125.

Chorchón, L., C. (2007), The Theory of Implementation: What did We Learn? Working Paper, Universitat Carlos III, Spain.

——— (1996), *The Theory of Implementation of Socially Optimal Decisions in Economics.* Palgrave McMillan, USA.

Clark, E. (1971), Multi-Part Pricing of Public Goods. *Public Choice* 11, 17-33.

Danilov, V., I. and Sotskov, A., I. (2002), *Social Choice Mechanisms.* Studies in Economic Design, Springer-Verlag, Germany.

Eliaz, K. (2002), Fault Tolerant Implementation. *Review of Economic Studies* 69, 589-610.

Farquharson, R. (1969), *Theory of Voting.* Yale University Press, New Haven.

Fishburn, P. C. (1973), *The Theory of Social Choice.* University Press, Princeton.

Groves, T. (1973), Incentive in Teams. *Econometrica* 41, 617-31.

Hayek, F. (1945), The Use of Knowledge in Society. *American Economic Review* 35, 519-530.

——— (1935), The Present State of the Debate. In *Collectivist Economic Planning*, edited by Hayek, F., George Routledge and Sons, London, 201-243.

Herrero, M., J. and Srivastava, S. (1992), Implementation via Backward Induction. *Journal of Economic Theory* 56, 70-88.

Hurwicz, L. (1960), Optimality and Informational Efficiency in Resource Allocation Processes. In *Mathematical Methods in the Social Sciences*, edited by K. J. Arrow, S. Karlin and P. Suppes, Stanford University Press, 27-46.

——— (1972), On Informationally Decentralized Systems. In *Decision and Organization*, edited by R. Radner and C. B. McGuire, North-Holland, Amsterdam, 297-336.

Jackson, M., O. and Wilkie, S: (2005), Endogenous Games and Mechanisms: Side Payments Among Players. *Review of Economic Studies* 72, 543-566.

Jackson, M., O. (2001), A Crash Course in Implementation Theory. *Social Choice and Welfare* 18, 655-708.

Jackson, M., O., Palfrey, T. and Srivastava, S. (1994), Undominated Nash Implementation in Bounded Mechanisms. *Games and Economic Behavior* 6, 474-501.

Jackson, M., O. (1992), Implementation in Undominated Strategies: A Look at Bounded Mechanisms. *Review of Economic Studies* 59, 757-775.

——— (1991), Bayesian Implementation. *Econometrica* 59, 461-477.

Kuhn, H. and Nasar, S. (2002), *The Essential John Nash*. Princeton University Press, Princeton.

Lange, O. (1937), On the Theory of Socialism, Part II. *Review of Economic Studies* 4, 123-142.

Lerner, A. (1936), A Note on Socialist Economies. *Review of Economic Studies* 4, 72-76.

Maskin, E. and Sjöström, T. (2002), Implementation Theory. Chapter 5 in *Handbook of Social Choice and Welfare*, edited by K. J. Arrow and A. K. Sen, Elsevier.

Matsushima, H. (1988), A New Approach to the Implementation Problem. *Journal of Economic Theory* 45, 128-144.

Milgrom, P. (1981), An Axiomatic Characterization of Common Knowledge. *Econometrica* 49, 219-222.

Mirrlees, J. (1971), An Exploration in the Theory of Optimal Income Taxation. *Review of Economic Studies* 38, 175-208.

Mises, L. von (1920), Die Wirtschaftsrechnung im Sozialistischen Gemeinwesen. In *Collectivist Economic Planning* (1935), edited by Hayek, F., George Routledge and Sons, London, 87-103.

Mookherjee, D. and Reichelstein, S. (1990), Implementation via Augmented Revelation Mechanisms. *Review of Economic Studies* 57, 453-476.

Moore, J. (1992), Implementation, Contracts, and Renegotiation in Environments with Complete Information. In *Advances in Economic Theory: Sixth World Congress*, vol. I, edited by J.-J. Laffont. Cambridge, UK, Cambridge University Press.

Moore, J. and Repullo, R. (1988), Subgame Perfect Implementation. *Econometrica* 56, 1191-1220.

Nurmi, H. (1987), *Comparing Voting Systems*. D. Reidel Pub. Co., Dordrecht.

Palfrey, T. (1992), Implementation in Bayesian Equilibrium: the Multiple Equilibrium Problem in Mechanism Design. In *Advances in Economic Theory: Sixth World Congress*, vol. I, edited by J.-J. Laffont. Cambridge, UK, Cambridge University Press.

Palfrey, T. and Srivastava, S. (1991), Nash-Implementation Using Undominated Strategies. *Econometrica* 59, 479-501.

Roberts, K. (1979), The Characterization of Implementable Choice Rules. In *Aggregation and Revelation of Preferences*, edited by J.-J. Laffont, North-Holland, Netherlands.

Rothschild, M. and Stigliz, J. (1976), Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect information. *Quarterly Journal of Economics* 90, 630-49.

Saari, D., G. (1994), *Geometry of Voting*. Springer-Verlag, Germany.

Samuelson, P. (1954), The Pure Theory of Public Expenditure. *Review of Economics and Statistics* 36, 387-389.

Selten, R. (1975), Re-Examination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory* 4, 25-55.

Sen, A. (1986), Social Choice Theory. In *Handbook of Mathematical Economics*, vol. 3, edited by K. J. Arrow and M. D. Intriligator. North-Holland, Amsterdam.

Serrano, R. (2003), The Theory of Implementation of Social Choice Rules. Working Paper No. 2003-19, Brown University, USA.

Sjöström, T. (1993), Implementation in Perfect Equilibrium. *Social Choice and Welfare* 10, 97-106.

Thomson, W. (1979), Maximin Strategies and Elicitation of Preferences. In: Laffont, J. J. (ed.): *Aggregation and Revelation of Preferences.* North-Holland, Amsterdam, 245-268.

Vartiainen, H. (2007), Subgame Perfect Implementation: A Full Characterization. *Journal of Economic Theory* 133, 111-126.

——— (1999), Studies in Mechanism Design and Game Theory. Academic Dissertation, University of Helsinki and RUSEG, mimeo.

Vickrey, W. (1961), Counterspeculation, Auctions and Competitive Sealed Tenders. *Journal of Finance* 16, 1-17.

Williams, S. (2008), *Communication in Mechanism Design: A Differential Approach.* Cambridge University Press, USA.

Part II

The Original Research Papers

ARTICLE 1

On the Communication Requirements of Secure and Dominant Strategy Implementation

On the Communication Requirements of Secure and Dominant Strategy Implementation

Ville Korpela * †

Public Choice Research Centre,

University of Turku

December 3, 2009
Last Revision: May 20, 2011

Abstract

The literature on communication requirements of goal functions (or social choice rules) mostly assume that agents act sincerely. In the case that incentive compatibility must be met, these results give only a lower bound for the number of messages needed. In this paper, we show that the lower bound of sincere behavior is also an upper bound for a wide class of goal functions, when the appropriate incentive compatibility requirement is secure or dominant strategy implementation.

JEL Classification: D71; D78; D82; D83
Keywords: Communication requirements; Decentralized mechanism; Direct revelation mechanism; Implementation; Informational efficiency

1. INTRODUCTION

When the designer of a social choice mechanism has no reason to assume that agents know anything about the environment but merely their own preferences, dominant strategy equilibrium is the appropriate solution con-

*E-mail: ville.korpela@utu.fi.

†Many thanks to Hannu Vartiainen for useful discussions. I Would also like to thank Klaus Kultti for comments on an earlier draft.

cept to be used in an implementation problem.[1] On the other hand, direct revelation mechanism is able to implement truthfully any goal function that is implementable in dominant strategies (Dasgupta et al., 1979), so that implementation problems look almost trivial under this informational assumption. Furthermore, even if truthful implementation is not considered satisfactory, as there may be other dominant strategy equilibria besides the truth-telling one, an additional property called weak non-bossiness is enough to guarantee that direct revelation mechanism has no bad equilibria (Saijo et al., 2007; Mizukami and Wakayama, 2007).

Unfortunately, this clean view does not address one important practical question: What if the direct revelation mechanism is informationally very inefficient? When mechanism design problem has m possible (social) outcomes, the number of conceivable linear preferences alone is $m!$. Therefore, the number of messages needed in a direct revelation mechanism can, at least potentially, be exponentially larger than the number of messages needed in a minimal indirect mechanism. Even though the Gibbard-Satterthwite theorem (Gibbard, 1973; Satterthwaite, 1975) [2] tells us that the preference domain cannot be this large for every agent (if the goal function is to be implementable at all), it can certainly be many times larger than the number of outcomes. This has been made evident by many interesting dominant strategy implementable goal functions recently presented in the literature: Generalized median voter rules in single-peaked environments (Moulin, 1980), serial cost sharing rules (Moulin and Shenker, 1992) and many others (Saijo et al., 2007; Barberà, 2001).

The main purpose of this paper is to find out how many messages are needed in a minimal indirect mechanism when the relevant incentive constraint is that of secure or dominant strategy implementation. Quite surprisingly, it

[1] Dominant strategy implementation can be interpreted as a sort of *robust implementation*. See Bergemann and Morris (2008), for example.

[2] If the preferences domain is unrestricted (i.e., universal) and the goal function is sovereign (i.e., onto), then strategy-proofness implies that the goal function must be dictatorial.

turns out that the minimal number of messages (or strategies) needed in an implementing game form is often exactly the same as the minimal number of messages needed in a mechanism that realizes the goal function in a decentralized way. The literature on communication requirements of goal functions has mostly concentrated on resource allocation processes (e.g., Hurwicz, 1972; Osana, 1987)[3] or completely neglected the incentive side (e.g., Mount and Reiter, 1974; Segal, 2007). There are few notable exceptions (e.g., Williams, 1984; Reichelstein and Reiter, 1988), but contrary to this paper, the relevant incentive constraint has been that of Nash equilibrium. Moreover, these papers usually assume some structure on the set of outcomes (manifold or a topological space), whereas we assume no abstract structure at all.

The most common objection raised against indirect mechanisms is that agents will have to learn more complicated rules, instead of simply announcing their own preferences, which they do not necessarily even understand. Consequently, it is not obvious whether anything can be accomplished using indirect mechanisms. Despite this, there are at least two things that suggest the question might be important. First, and foremost, the direct revelation mechanism is mostly just a representational tool. In the complicated real world institutional environment, strategies are not usually preferences themselves (even though they can be). It may then be that the role of complexity and informational efficiency are central in explaining the evolution of these institutions. These are, in fact, some of the questions that we do not yet understand well.[4] Second, sometimes it may be preferable to use indirect mechanisms because then the agents do not need to know their preferences accurately. It may be demanding for an agent to know her preferences in a fine tuned manner, and indirect mechanisms can have one strategy that is dominant for an entire class of preferences. Think about the median voter

[3] Mainly the competitive price system.

[4] There is no consensus about the definition of an institution. Are instituitons part of the equilibrium concept, or are they rules of the game. See Hurwicz (1994) and Oström (1986), for example.

rule, for example.[5] If the number of voters is odd and the domain consists of all possible single-peaked preferences, then this rule is dominant strategy implementable. But the outcome only depends on some features of the preferences, namely the "peak". So the rule would be dominant startegy implementable even if every agent is asked to announce only the "peak". That is, to vote for the most favored candidate.

The rest of the paper is organized as follows. In section 2, some properties that are needed in characterizing implementable goal functions are given. After this we introduce a method, presented in Hurwicz and Reiter (2008), which can be used to construct a decentralized mechanism with minimal number of messages. Since every game form that implements in Nash or dominant strategy equilibrium is a decentralized mechanism, it is conceivable that this minimal decentralized mechanism can be expanded to obtain a minimal implementing game form.[6] In section 3, we prove that this can indeed be done. More than this, it turns out that the minimal decentralized mechanism of Hurwicz and Reiter (2008) is quite generally also incentive compatible. Finally, section 4 concludes the paper. It is not possible to go through all set theoretic concepts that are needed, so a fair amount of familiarity is assumed.

2. DEFINITIONS, PRELIMINARIES AND NOTATIONAL CONVENTIONS

We denote the set of *outcomes* by Z and the set of *agents* by $N = \{1, \ldots, n\}$. A typical element of N is denoted by i or j, and a typical element of Z is denoted by x, y or z, and so forth. We assume that there are at least two agents $n \geq 2$. Every agent i is endowed with a complete and transitive preference ordering θ_i over the outcomes Z and the set of all admissible preferences orderings for agent i is denoted by Θ_i. The strict part of θ_i is denoted by $P(\theta_i)$ and the indifference part by $I(\theta_i)$. Then, using the notation $\theta = (\theta_1, \ldots, \theta_n) \in \Theta \equiv \times_{i \in N} \Theta_i$, a *goal function* is any function

[5] See Moulin (1980).

[6] See Williams (1984) for a rigorous definition of an embedding problem in the case of Nash equilibrium.

$f : \Theta \to Z$ that associates a unique alternative $f(\theta) \in Z$ with every profile of preference orderings $\theta \in \Theta$. A few properties of goal functions will be needed.

Definition 1 (*Strategy-Proofness*). A goal function f satisfies *strategy-proofness* if, for all $\theta \in \Theta$ and all $i \in N$, there is no $\theta'_i \in \Theta_i$, such that

$$f(\theta'_i, \theta_{-i}) P(\theta_i) f(\theta).$$

Strategy-proofness is a strong incentive compatibility condition. It says that no agent can gain by lying regardless of whether everyone else is speaking the truth or not.[7]

Definition 2 (*Rectangular Property*). A goal function f satisfies *rectangular property* if, for all $\theta, \theta' \in \Theta$, we have

$$f(\theta') I(\theta_i) f(\theta_i, \theta'_{-i}) \text{ for all } i \in N \quad \Rightarrow \quad f(\theta') = f(\theta).$$

Rectangular property is a technical rule that emerges as a necessary condition for secure implementation (Saijo et al., 2007) defined later on.

Definition 3 (*Weak Non-Bossiness*).[8] A goal function f satisfies *weak non-bossiness* if, for all $\theta \in \Theta$, all $i \in N$, and all $\theta'_i \in \Theta_i$, we have

$$f(\theta_i, \theta_{-i}) \neq f(\theta'_i, \theta_{-i}) \quad \Rightarrow \quad \exists\, \theta^*_{-i} \in \Theta_{-i} : \; \neg\, f(\theta_i, \theta^*_{-i}) I(\theta_i) f(\theta'_i, \theta^*_{-i}).$$

Basically, weak non-bossiness requires that whenever an agent can unilaterally change the outcome, this must affect her own utility, at least in some cases.

A *mechanism*, as defined by Hurwicz and Reiter (2008), is a triplet $\pi = (\mu, M, h)$, where M is the (common) *message space*, $\mu : \Theta \to M$ is the (group) *equilibrium message correspondence* and $h : M \to Z$ is the *outcome function.* We say that mechanism π *realizes* f if

$$h \circ \mu(\theta) = f(\theta) \text{ for all } \theta \in \Theta, \tag{1}$$

[7] See Barberà (2001), for example, for more on strategy-proofness.

[8] This condition is called *Quasi-Strong-Non-Bossiness* in Mizukami and Wakayama (2007). A similar condition has been presented in Satterthwaite and Sonnenschein (1981).

and call π *informationally decentralized*, or simply *decentralized*, if for every agent i there exists a correspondence $\mu_i : \Theta_i \to M$, such that

$$\mu(\theta) = \bigcap_{i \in N} \mu_i(\theta_i) \quad \text{for all} \quad \theta \in \Theta.^{9} \tag{2}$$

This condition states that equilibrium can be obtained in a *privacy-preserving* way. That is, the message sent by an agent depends on her own preferences only. One should notice that any goal function f can be realized with the *decentralized direct revelation mechanism*, defined by $M = \Theta$, $\mu_i(\theta_i) = \{\theta' \in \Theta \mid \theta'_i = \theta_i\}$ and $h(\theta) = f(\theta)$. Therefore, the question is not whether a given goal function f can be realized, but rather, how can it be realized with the minimal number of messages $|M|$? This has been solved in Hurwicz and Reiter (2008). Below we explain the constructive method in its essentials.

By a *covering* of Θ, we mean a class $C = \{K \mid K \subseteq \Theta\} \subseteq 2^{\Theta}$, such that $\bigcup_{K \in C} K = \Theta$. Covering C is called *f contour contained* (in brief, f-cc) if for all $K \in C$ there exists $z \in Z$, such that $K \subseteq f^{-1}(z)$, and it is called *rectangular* if all sets $K \in C$ have the structure of a cartesian product. Any f-cc and rectangular covering C, that does not contain *redundancy*,[10] has a *system of distinct representatives* (SDR),[11] that is, a function $\Lambda : C \to \Theta$ that satisfies two properties: (1) $\Lambda(K) \in K$ for all $K \in C$, and (2) $K' \neq K''$ implies $\Lambda(K') \neq \Lambda(K'')$.

For any covering C of Θ, define a correspondence $\Omega : \Theta \to C$, $\Omega(\theta) = \{K \in C \mid \theta \in K\}$ and let $\upsilon : \Lambda(C) \to M$ be a bijective (onto) mapping.[12]

---

[9] Let us assume that every agent i acts according to some message verification protocol $\boldsymbol{m}_i^{t+1} = g_i(\boldsymbol{m}^t, \theta_i)$, where $\boldsymbol{m}^t$ is the vector of messages sent at time t and $\boldsymbol{m}_i^{t+1}$ is the reply of agent i at time $t+1$. With this interpretation, μ can be thought of as representing a static equilibrium of a message verification scenario *i.e.*

$$\mu(\theta) = \{\boldsymbol{m} \in M \mid \boldsymbol{m}_i = g_i(\boldsymbol{m}, \theta_i) \text{ for all } i \in N\} = \bigcap_{i \in N} \{\boldsymbol{m} \in M \mid \boldsymbol{m}_i = g_i(\boldsymbol{m}, \theta_i)\} = \bigcap_{i \in N} \mu_i(\theta_i).$$

For more details see Hurwicz (1994).

[10] Formally, we do not have $K \subseteq \bigcup_{K' \in C \setminus K} K'$ for any $K \in C$.

[11] A basic theorem on systems of distinct representatives is proven in Hall (1948).

[12] We could simply choose a subset of Θ as our message space. The main purpose of

Theorem 1 (*Hurwicz and Reiter, 2008*). Let C be an f-cc and rectangular covering of Θ that does not contain redundancy. The mechanism $\pi_C = (\mu, M, h)$, defined by the following two conditions:

(i) $\mu = v \circ \Lambda \circ \Omega$, and

(ii) $h = f \circ v^{-1}$,

will realize f. ■

To complete the description, we need to address two more questions. First, can μ be decentralized? Second, what are the properties of C that guarantee the minimality of M? To this end, we define for every agent $i \in N$, a correspondence $\Omega_i : \Theta_i \to C$, $\Omega_i(\theta_i) = \{K \in C \mid \theta_i \in proj_i K\}$.[13]

Theorem 2 (*Hurwicz and Reiter, 2008*). The mechanism π_C satisfies condition (2) when we choose $\mu_i = v \circ \Lambda \circ \Omega_i$ for all $i \in N$. Hence, mechanism π_C can indeed be decentralized. ■

To answer the second question, we must emphasize that any decentralized mechanism π can be produced with the previous method by using some rectangular and f-cc covering of Θ. That is, for any decentralized mechanism π, there exists a rectangular and f-cc covering C of Θ, such that $\pi = \pi_C$. Consequently, it should be obvious that the minimal size of the message space M is somehow connected with C being a maximally coarse covering.[14] Unfortunately, the relation "*coarsening*" is not complete. There can exist, and usually do exist, many maximally coarse coverings with different or equal numbers of sets.

Luckily, we do not need to dwell on the question of constructing a maximally

coding function v is to make it explicit that agents do not have to transmit a complete characterization of their preferences, only an abstract message.

[13]The set $proj_i K$ is formally defined as $\{\theta_i \in \Theta_i \mid (\theta_i, \theta_{-i}) \in K$ for some $\theta_{-i} \in \Theta_{-i}\}$.

[14]Covering C' is a *coarsening* of C if for every $K \in C$ there is $K' \in C'$, such that $K \subseteq K'$. It is a *proper coarsening* if it is a coarsening and $K \subset K'$ for some $K \in C$ and $K' \in C'$. A *maximally coarse covering* can then be defined as a covering that does not have any proper coarsenings. Notice that a maximally coarse, f-cc and rectangular covering cannot contain redundancy (see Theorems 1 and 2).

coarse covering with the minimal number of sets. It turns out that when the goal function f is sufficiently "*well-behaved*" (e.g., implementable), then the construction of Hurwicz and Reiter (2008), reproduced in Theorems 1 and 2, can be used to build an incentive compatible mechanism with a minimal message space starting from *any* maximally coarse covering of Θ. [15]

Before presenting the main results, we need a few definitions to make a clear-cut distinction between realization and an incentive compatible realization. A *Game form* is a tuple $G = (S, g)$, where $S = S_1 \times \cdots \times S_n$ is the *strategy space*, the set S_i is a strategy space of agent i and $g : S \to Z$ is the *outcome function.* For a fixed preference profile $\theta \in \Theta$, this game form defines a game $\Gamma(\theta) = (G, \theta)$ in normal form. Strategy profile $\boldsymbol{s}^*$ is a *Nash equilibrium* of the game $\Gamma(\theta)$ if $g(\boldsymbol{s}^*)\theta_i g(s_i, \boldsymbol{s}^*_{-i})$ for all $i \in N$ and all $s_i \in S_i$. The set of all Nash equilibria in the game $\Gamma(\theta)$ is denoted by $NE\big[\Gamma(\theta)\big]$. We say that game form G *implements* f *in Nash equilibrium* if

$$g\left(NE\big[\Gamma(\theta)\big]\right) = f(\theta) \;\text{ for all }\; \theta \in \Theta. \tag{3}$$

In a similar way, strategy profile $\boldsymbol{s}^*$ is a *dominant strategy equilibrium* of the game $\Gamma(\theta)$ if $g(\boldsymbol{s}^*_i, \boldsymbol{s}_{-i})\theta_i g(\boldsymbol{s})$ for all $i \in N$ and all $\boldsymbol{s} \in S$. The set of all dominant strategy equilibria in the game $\Gamma(\theta)$ is denoted by $DOM\big[\Gamma(\theta)\big]$, and a game form G *implements* f *in dominant strategies* if equation (3) holds when $NE = DOM$. If game form G implements f in Nash and dominant strategy equilibrium, then according to Saijo et al. (2007), we call it *securely implementable.*[16]

[15] Hurwicz and Reiter (2008) advance two methods of constructing a maximally coarse covering. The simpler one, called *reflexive rectangular method* (rRM), works heuristically as follows: Take an arbitrary preference profile $\theta \in \Theta$ and form a largest possible rectangle contained in the same contour set (outcome is not unique). Choose a new preference profile that is outside the class already formed and continue until a covering of Θ is obtained. The end result will always be a maximally coarse covering in the set of all rectangular and f-cc coverings (but not generally the minimal).

[16] Double implementation, a term that was coined by Maskin (1979), is called secure implementation if the two solution concepts are Nash and dominant strategy equilibrium. To better appreciate this implementation form, see Repullo (1985) and Saijo et al. (2007).

3. MAIN RESULTS

Even though the problem of finding an informationally efficient decentralized mechanism and the problem of finding a minimal implementing game form are fundamentally very different, we can still use the former as a starting point when seeking an answer to the latter. If we want to implement goal function f using a solution concept that is decentralizable,[17] then every game form $G = (S, g)$ that implements f can be used to form a decentralized mechanism with a (common) message space of cardinality $|S|$. It is then conceivable, but by no means *a priori* certain, that the minimal decentralized mechanism can be used to obtain a minimal implementing game form. The main purpose of this section is to show that this can be done – not only that, but fairly easily.

Next, we need to complete the construction given in Theorems 1 and 2 to show how the individual message space M_i can be recovered. Let f be a goal function and C an f-cc and rectangular covering of Θ that is maximally coarse (and hence does not contain redundancy). Moreover, for all agents $i \in N$, let $C[i]$ be the maximally coarse covering of Θ_i that satisfies the following condition: For all $A \in C[i]$ and all $K \in C$,

$$\text{either} \quad A \subseteq proj_i K \quad \text{or} \quad A \cap proj_i K = \emptyset. \tag{4}$$

Now, let $v_i : C[i] \to M_i$ be a bijective (coding) function for all $i \in N$, and define a decentralized mechanism $\widehat{\pi}_C = (\mu, M, h)$ by the following three rules:[18]

$$(i)\ M = M_1 \times \cdots \times M_n \tag{5}$$

$$(ii)\ \mu(\theta) = \bigcap_{i \in N} \mu_i(\theta_i) \text{ and } \mu_i(\theta_i) = \{\boldsymbol{m} \in M \mid \theta_i \in v_i^{-1}(\boldsymbol{m}_i)\} \tag{6}$$

$$(iii)\ h(\boldsymbol{m}) = f\left(\mathop{\times}_{i \in N} v_i^{-1}(\boldsymbol{m}_i) \right) \tag{7}$$

[17] The Nash equilibrium correspondence $f : \Theta \to NE\big[\Gamma(\theta)\big]$ of game form $G = (S, g)$ can be decentralized through the mechanism $\pi = (\mu, S, g)$, where $\mu_i(\theta_i) = \{\boldsymbol{s} \in S \mid g(\boldsymbol{s})\theta_i g(s', \boldsymbol{s}_{-i}) \text{ for all } s' \in S_i\}$. Not all equilibrium concepts are decentralizable in this fashion, such as strong equilibrium, for example.

[18] For any $K \subseteq \Theta$, we denote $f(K) = \{f(\theta) \mid \theta \in K\}$ as usual.

The general idea behind this construction is sketched in Figure 1 below. Here, $C = \{K_i \mid i = 1, 2, 3, 4, 5 \text{ or } 6\}$ and $M_i = \{m_1^i, m_2^i, m_3^i, m_4^i\}$ for agent $i \in \{1, 2\}$.

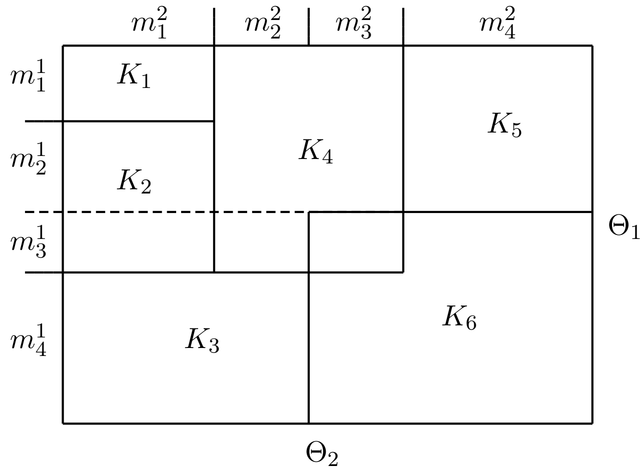

FIGURE 1. Constructing the message space of mechanism $\widehat{\pi}_C$

The following important Lemma makes the construction in equations (4)-(7) more transparent.

Lemma. Mechanism $\widehat{\pi}_C$ is well-defined and it realizes f.

Proof. Notice that a maximally coarse covering $C[i]$ of Θ_i must exist due to the fact that Θ_i itself satisfies the defining condition (4). We shall first show that $C[i]$ is unique. Suppose there are two different maximally coarse coverings of Θ_i, $C[i]$ and $C^*[i]$, which both satisfy (4). Then, for some $A \in C[i]$ and $A', A'' \in C^*[i]$, such that $A' \neq A''$, we have

$$A \cap A' \neq \emptyset \ \text{ and } \ A \cap A'' \neq \emptyset.$$

But then, the class $(C^*[i] \setminus \{A', A''\}) \cup (A' \cup A'')$ also satisfies condition (4),[19] which is a contradiction with the fact that $C^*[i]$ is maximally coarse. For $\widehat{\pi}_C$ to be well-defined, we still have to show that $\underset{i \in N}{\times} v_i^{-1}(\boldsymbol{m}_i)$ is included in the same contour set for all $\boldsymbol{m} \in M$. Choose any $\boldsymbol{m} \in M$ and $\theta \in \underset{i \in N}{\times} v_i^{-1}(\boldsymbol{m}_i)$. Let $K \in C$ be such that $\theta \in K$, so that by definition, we

[19]Since $A \cup A'$ and $A \cup A''$ must satisfy condition (4), as $\theta \in A \cap A'$ and $\psi \in A \cap A''$ for some $\theta, \psi \in \Theta$, $A' \cup A''$ must also satisfy it.

have $v_i^{-1}(\boldsymbol{m}_i) \subseteq proj_i K$ for all $i \in N$. Since C is rectangular, this implies $\underset{i\in N}{\times} v_i^{-1}(\boldsymbol{m}_i) \subseteq K$, and since C is f-cc, the set $\underset{i\in N}{\times} v_i^{-1}(\boldsymbol{m}_i)$ must be included in the same contour set, and hence the outcome function h is well-defined. The fact that $\widehat{\pi}_C$ realizes f can be verified by a direct computation:

$$h(\mu(\theta)) = h\left(\bigcap_{i\in N} \mu_i(\theta_i)\right) = h\left(\{\boldsymbol{m} \in M \mid \theta \in \underset{i\in N}{\times} v_i^{-1}(\boldsymbol{m}_i)\}\right) = f(\theta),$$

where the last equality must hold, since h is well-defined. ■

The construction given in equations (4)-(7) is important. Hurwicz and Reiter (2008) has shown that the mechanism $\widehat{\pi}_C$ has a minimal number of messages $|M_i|$, in the set of all decentralized mechanisms that realize f, when C is chosen as a maximally coarse covering with a minimal number of sets. In fact, as we shall see shortly, this construction even preserves incentive compatibility. The following example illustrates the theoretical construction in equations (4)-(7) and gives a preview of the theorems to come.

Example 1. Let $N = \{1, 2\}$, $Z = \{x, y, z\}$ and $\Theta = \{\theta_1, \theta_2, \theta_3\} \times \{\psi_1, \psi_2, \psi_3\}$. The preferences are given in Table 1 below and goal function f is defined in Figure 2 below.

TABLE 1

θ_1	θ_2	θ_3	ψ_1	ψ_2	ψ_3
x	x	z	x	z	x
y	y, z	x, y	y	y	y, z
z			z	x	

It is straightforward to show that the direct revelation mechanism $G = (\Theta, f)$ implements f in dominant strategies. It satisfies strategy-proofness and weak non-bossiness (see Theorem 6). The unique maximally coarse, rectangular and f-cc covering of Θ is $C = \{K_1, K_2, K_3, K_4\}$, where

$$K_1 = \{\theta_1, \theta_2\} \times \{\psi_1, \psi_3\},\ K_2 = \{\theta_1, \theta_2\} \times \{\psi_2\},$$

$$K_3 = \{\theta_3\} \times \{\psi_1, \psi_3\} \text{ and } K_4 = \{\theta_3\} \times \{\psi_2\}.$$

		Agent 2		
		ψ_1	ψ_2	ψ_3
	θ_1	x	y	x
Agent 1	θ_2	x	y	x
	θ_3	y	z	y

FIGURE 2. A diagram of goal function f

Hence, the maximally coarse covering $C[1]$ of Θ_1 is $\{\{\theta_1, \theta_2\}, \{\theta_3\}\}$ and the maximally coarse covering $C[2]$ of Θ_2 is $\{\{\psi_1, \psi_3\}, \{\psi_2\}\}$. Using the construction in equations (4)-(7) gives us the mechanism in Figure 3 below.

		Agent 2	
		m_1^2	m_2^2
Agent 1	m_1^1	x	y
	m_2^1	y	z

FIGURE 3. The mechanism given by equations (4)-(7)

Again, it is straightforward to show that the game form defined by this mechanism implements f in dominant strategies.[20] This must be the game form with minimal number of strategies, since f cannot be even realized in a decentralized way with less than $|M_1| = |M_2| = 2$ messages. □

[20]A game form G_π defined by the mechanism $\pi = (\mu, M, h)$ is $G_\pi = (M, h)$. That is, the message verification protocol μ_i can be freely chosen by agent i.

In the remainder of this paper we show that the phenomenon in Example 1 is quite general.

3.1 The Case of Secure Implementation

The following theorem presents a necessary and sufficient condition for a goal function f to be securely implementable.

Theorem 3 (*Saijo et al., 2007*). A goal function f is securely implementable if and only if both strategy-proofness and rectangular property hold. ■

The proof of this theorem is simple and elegant. If the two properties hold, then the direct revelation mechanism $G = (\Theta, f)$ implements f.[21] Still, the strategy space $S = \Theta$ of this game form can be informationally very inefficient, which makes it important to find out how many strategies are needed in a minimal indirect game form implementing f. The following theorem will answer this question.

Theorem 4. Let f be securely implementable and $\widehat{\pi}_C = (\mu, M, h)$, the mechanism defined by equations (4)-(7) using a *maximally coarse*, *rectangular* and *f-cc* covering C. The game form $G = (S, g) = (M, h)$ implements f securely.

Proof. Denote $\mathit{\Gamma}(\theta) = (G, \theta)$. To prove this theorem, we have to verify two things: For every preference profile $\theta \in \Theta$, there exists a strategy profile $\boldsymbol{m} \in DOM\big[\mathit{\Gamma}(\theta)\big]$ such that $h(\boldsymbol{m}) = f(\theta)$, and for every strategy profile $\boldsymbol{m} \in NE\big[\mathit{\Gamma}(\theta)\big]$ we have $h(\boldsymbol{m}) = f(\theta)$. First, choose any $\theta \in \Theta$ and let $\boldsymbol{m}^d \in M$ be such that $\theta \in \underset{i \in N}{\times} v_i^{-1}(\boldsymbol{m}_i^d)$. We show that $\boldsymbol{m}^d \in DOM\big[\mathit{\Gamma}(\theta)\big]$. Assume the contrary. That is, there exists a message profile $\boldsymbol{m}' \in M$, such that $h(\boldsymbol{m}')P(\theta_j)h(\boldsymbol{m}_j^d, \boldsymbol{m}'_{-j})$ for some $j \in N$. Since $\widehat{\pi}_C$ realizes f in a decentralized way, we have $h(\boldsymbol{m}') = f(\theta')$ and $h(\boldsymbol{m}_j^d, \boldsymbol{m}'_{-j}) = f(\theta_j, \theta'_{-j})$ for all $\theta' \in \underset{i \in N}{\times} v_i^{-1}(\boldsymbol{m}'_i)$, so that $f(\theta')P(\theta_j)f(\theta_j, \theta'_{-j})$ must hold for some

[21] This is the hard direction in all implementation proofs, since some kind of *canonical mechanism* has to be found.

$\theta'_{-j} \in \Theta_{-j}$[22]. This is a contradiction with the fact that f is strategy-proof, and hence $\boldsymbol{m}^d \in DOM\big[\Gamma(\theta)\big]$.

To verify the second part, let us assume that $\boldsymbol{m} \in NE\big[\Gamma(\theta)\big]$. Since the strategy $\boldsymbol{m}_i^d$ defined in the previous paragraph is a dominant strategy for every agent $i \in N$, we must have $h(\boldsymbol{m}_i^d, \boldsymbol{m}_{-i})\theta_i h(\boldsymbol{m})$ for all $i \in N$. As $\boldsymbol{m}$ is Nash equilibrium under θ, this implies that $h(\boldsymbol{m}_i^d, \boldsymbol{m}_{-i})I(\theta_i)h(\boldsymbol{m})$ for all $i \in N$. Thus, since $\widehat{\pi}_C$ realizes f in a decentralized way, there must exist $\theta' \in \underset{i\in N}{\times} v_i^{-1}(\boldsymbol{m}_i)$ such that $f(\theta_i, \theta'_{-i})I(\theta_i)f(\theta'_i, \theta'_{-i})$ for all $i \in N$. Now, the fact that f satisfies rectangular property implies $f(\theta') = f(\theta)$, and since $\theta' \in \underset{i\in N}{\times} v_i^{-1}(\boldsymbol{m}_i)$, we finally get $h(\boldsymbol{m}) = f(\theta') = f(\theta)$, as required. ■

This theorem implies the following strong result as a direct corollary.

Corollary 1. The minimal number of messages needed for a decentralized realization of a securely implementable goal function is exactly the same as the minimal number of strategies needed to implement it.

Proof. The number of strategies needed in implementation is at least as great as the number of messages needed to realize in a decentralized way. Hence, the result follows from Theorem 4. ■

Next, we prove this result directly to generate deeper insight into equations (4)-(7).

Theorem 5. Let f be securely implementable and $\widehat{\pi}_C = (\mu, M, h)$, the mechanism defined in Theorem 4. Then, f cannot be securely implemented with less than $\sum_{i=1}^n |M_i|$ strategies (or less that $|M|$ strategy profiles).

Proof. We prove this claim by showing that every agent $i \in N$ needs at least $|M_i|$ strategies. For the sake of contradiction, suppose that agent i could have fewer strategies in a securely implementing game form $G = (S, g)$. Then, by the definition of message space M, there must exist $K, K' \in C$, $K \neq K'$, such that the same strategy $s^d \in S_i$ is dominant for some $\theta_i \in proj_i K \setminus proj_i K'$ and $\theta'_i \in proj_i K'$. Choose $\theta'_{-i} \in \Theta_{-i}$ in such way that

[22]Notice that $(\theta_i, \theta'_{-i}) \in v_i^{-1}(\boldsymbol{m}_i^d) \times \left(\underset{j\in N\setminus\{i\}}{\times} v_j^{-1}(\boldsymbol{m}'_j) \right)$.

$(\theta'_i, \theta'_{-i}) \in K'$ and let $\boldsymbol{s}^d_{-i}$ be a vector of dominant startegies for θ'_{-i}. Since dominant strategies only depend on the agent's own preferences, it must be that $(s^d_i, \boldsymbol{s}^d_{-i}) \in DOM\big[\Gamma(\theta_i, \theta'_{-i})\big]$ and $(s^d_i, \boldsymbol{s}^d_{-i}) \in DOM\big[\Gamma(\theta'_i, \theta'_{-i})\big]$, which implies that $f(\theta_i, \theta'_{-i}) = f(\theta'_i, \theta'_{-i})$. As this must hold for any $\theta'_{-i} \in \underset{j \in N\setminus\{i\}}{\times} proj_j K'$, the covering

$$(C \setminus K') \cup \left(\{\theta_i \cup proj_i K'\} \underset{j \in N\setminus\{i\}}{\times} proj_j K' \right)$$

is f-cc and rectangular – a contradiction with the assumption that C is maximally coarse. ■

Theorems 4 and 5 do not seem to depend on the covering, as long as it is maximally coarse. We could easily show that this is due to the fact that maximally coarse covering is a unique partition when f is securely implementable.

Notice that any securely implementable goal function can be implemented with the direct revelation mechanism, which means that only outcomes in the range are needed in an implementing game form. That is, outcomes in $Z \setminus f(\Theta)$ are never needed. We can interpret this by saying that secure implementation is context independent, which is in stark contrast with Nash implementation[23] and dominant strategy implementation (Example 2 below). This is the reason that allows us to use the method of Hurwicz and Reiter (2008) as a starting point in the first place.

Example 2. Assume that $Z = \{x, y, z\}$ and $\Theta = \{\theta_1, \theta'_1\} \times \{\theta_2\}$. Define goal function f by the rule $f(\theta_1, \theta_2) = x$ and $f(\theta'_1, \theta_2) = y$, so that $z \notin f(\Theta)$. Moreover, define the preferences θ_1, θ'_1 and θ_2 by setting

$$xI(\theta_1)y, \; xI(\theta'_1)y \; \text{ and } \; xI(\theta_2)y.$$

It is obvious that f cannot be implemented in dominant strategies using alternatives only from $f(\Theta) = \{x, y\}$, since both agents would be indifferent about playing any strategy. Still, it can be implemented using $Z = \{x, y, z\}$,

[23]See Williams (1984), for example.

at least if the following preference reversals hold:

$$zP(\theta_1)x,\ xP(\theta_1')z \text{ and } xP(\theta_2)z.$$

Assuming that this is indeed the case, the game form in Figure 4 below will implement f. As a consequence of Theorem 3, f cannot be securely implemented even when using the whole set Z. For example, the game form in Figure 4 has a bad Nash equilibrium (s_1^2, s_2^2) for the preference profiles (θ_1', θ_2).

		Agent 2	
		s_2^1	s_2^2
Agent 1	s_1^1	x	z
	s_1^2	y	x

FIGURE 4. A game form implementing f in dominant strategies

□

There is nothing trivial in this example. The fact that all outcomes are indifferent for the agents involved, does not mean that they are indifferent from the social point of view (or from the principal point of view). Nonetheless, it is not completely harmless to use game forms that do not give outcomes from the range of a goal function if out of equilibrium strategies are played. As explained in Maskin and Moore (1999), this will inevitably raise some questions of credibility. If the designer cannot fully commit to implement anything that might arise as an outcome of the game, then agents can use this to initiate a renegotiation process by playing non-equilibrium strategies. It is then natural to assume that the outcome played before the renegotiation serves as an outside option. This will completely change the nature of the game and a new implementation concept is needed. It would be a lot harder to find a minimal implementing game form in this kind of setting, since more elaborate schemes could be used. Leaving this problem behind, Example 2 makes it very clear that Corollary 1 cannot hold for dominant strategy equilibrium as such.

3.2 The Case of Dominant Strategy Implementation

When the solution concept is dominant strategy equilibrium, we do not necessarily get a minimal implementing game form by applying equations (4)-(7) with maximally coarse covering. Some new strategies may have to be added. To proceed then, we need to know when our goal function can be implemented in dominant strategies by the direct revelation mechanism.

Theorem 6 (*Saijo et al., 2007*). The direct revelation mechanism $G = (\Theta, f)$ implements f in dominant strategies if and only if both strategy-proofness and weak non-bossiness holds. ■

It turns out that Theorems 4 and 5 hold for dominant strategies in the case described by Theorem 6.

Theorem 7. Let f satisfy strategy-proofness and weak non-bossiness and let $\widehat{\pi}_C = (\mu, M, h)$ be the mechanism defined in Theorem 4. The game form $G = (S, g) = (M, h)$ implements f in dominant strategies, and furthermore, f cannot be implemented with less than $\sum_{i=1}^{n} |M_i|$ strategies (or less that $|M|$ strategy profiles).

Proof. We already know that for any $\theta \in \Theta$, the messages profile $\boldsymbol{m}^d \in M$ that was defined in Theorem 4 is a dominant strategy equilibrium and $h(\boldsymbol{m}^d) = f(\theta)$. Only strategy-proofness was used to prove this. Now assume that $\boldsymbol{m}$ is any other dominant strategy equilibrium for θ and let $\theta' \in \underset{i \in N}{\times} v_i^{-1}(\boldsymbol{m}_i)$. Notice that, again by Theorem 4, $\boldsymbol{m}$ is a dominant strategy equilibrium also for θ'. By the definition of dominant strategy equilibrium, we must then have $h(\boldsymbol{m}_i, \boldsymbol{m}'_{-i}) I(\theta_i) h(\boldsymbol{m}_i^d, \boldsymbol{m}'_{-i})$ for all $i \in N$ and all $\boldsymbol{m}'_{-i} \in M_{-i}$, so that by the definition of $\widehat{\pi}_C$, we must also have $f(\theta'_i, \theta''_i) I(\theta_i) f(\theta_i, \theta''_i)$ for all $i \in N$ and all $\theta''_i \in \Theta_{-i}$. Consequently, weak non-bossiness implies $f(\theta_i, \theta_{-i}) = f(\theta'_i, \theta_{-i})$, and by the defintion of $\widehat{\pi}_C$ we then have $h(\boldsymbol{m}^d) = h(\boldsymbol{m}_i, \boldsymbol{m}_{-i}^d)$. But also $(\boldsymbol{m}_1, \boldsymbol{m}_{-1}^d)$ must be a dominant strategy equilibrium for θ (since $\boldsymbol{m}_1$ is a dominant strategy for θ_1). We can then replace $\boldsymbol{m}^d$ with $(\boldsymbol{m}_1, \boldsymbol{m}^d_{-1})$ in the previous argument and continue inductively in the number of agents to finally obtain $h(\boldsymbol{m}) = h(\boldsymbol{m}^d) = f(\theta)$ as required.

The proof that f cannot be implemented with less than $\sum_{i=1}^n |M_i|$ strategies is analogous to Theorem 5. ■

Corollary 1 also generalizes in an obvious manner.

Corollary 2. Assume that f satisfies strategy-proofness and weak non-bossiness. The minimal number of messages needed for a decentralized realization of f is exactly the same as the minimal number of strategies needed to implement it in dominant strategies. ■

To conclude the paper, let us take a look at one well-known example from the literature.

Example 3 (*The Generalized Median Voter Rule in a Single-Peaked Environment*).[24] For all $i \in N$, let $\Theta_i = \mathrm{U}_i$ be the set of all single-peaked preferences on some interval $[a, b] \subseteq \mathbb{R}_+$. The "peak" of the preference $u_i \in \mathrm{U}_i$ is denoted by $P(u_i)$. It is well-known that in this environment, the *generalized median voter rule*

$$f(u_1, \dots, u_n) = \text{median of } \big\{P(u_1), \dots, P(u_n), \alpha_1, \dots, \alpha_k\big\},$$

where n is odd, $k < n$ and the numbers $\alpha_j \in [a, b]$ are "*votes*" of the society,[25] is implementable in dominant strategies by the direct revelation mechanism. It is obvious that the strategy space of the minimal implementing game form, in this *unrestricted* preference domain, should consist of announcing only the "peak." This is exactly what our construction suggests. Even though the maximally coarse coverings C of Θ is a bit complicated to obtain, the only maximally coarse covering $C[i]$ that satisfies condition (4) must be

$$C[i] = \{A \subseteq \mathrm{U}_i \mid u_i, u_j \in A \text{ if and only if } P(u_i) = P(u_j)\}.$$

[24] We do not define all these, by now, well-known concepts here. See Moulin (1980,1983) or Saijo et al. (2007) for rigorous definitions. Moulin (1980) attributes the idea of single-peaked environments dating at least as far as Dummett and Farquharson (1961).

[25] Numbers $\alpha_1, \dots, \alpha_k$ are sometimes called "*phantom*" votes.

Note that this need not hold if the preference domain is restricted, but it is rather hard to say anything about this more general case without arbitrarily fixing the domain. □

4. CONCLUDING REMARKS

We have shown that the systematic procedure of Hurwicz and Reiter (2008) for constructing a minimal decentralized mechanism quite often preserves incentive compatibility. This is always so for a securely implementable goal function, but for a goal function implementable in dominant strategies, an additional property called weak non-bossiness in Saijo et al. (2007) is needed. A simple example from Mizukami and Wakayama (2007) verifies that the assumption of weak non-bossiness is essential. Let $N = \{1, 2\}$, $Z = \{x, y, z\}$ and $\Theta = \{\theta_1, \theta_1'\} \times \{\theta_2, \theta_2'\}$. The preferences are defined in the following way:

$$zP(\theta_1)xP(\theta_1)y, \; xP(\theta_1')zP(\theta_1')y, \; zI(\theta_2)xP(\theta_2)y \text{ and } yP(\theta_2')zI(\theta_2')x.$$

Goal function f is defined in the following way:

$$f(\theta_1, \theta_2) = z \text{ and } f(\theta_1, \theta_2') = f(\theta_1', \theta_2) = f(\theta_1', \theta_2') = x.$$

This goal function is implementable in dominant strategies. The indirect game form in Figure 5 below verifies this. The only dominant strategy for θ_1 is s_1^1; the only dominant strategy for θ_1' is s_1^2; the only dominant strategy for θ_2 is s_2^1; and the only dominant strategy for θ_2' is s_2^2. It is easy to verify that the outcome in these equilibria coincide with f. Since the direct revelation mechanism $G = (\Theta, f)$ does not implement f,[26] and on the other hand, the direct revelation mechanism must be inside any game form that implements f, the game form in Figure 5 must have a minimal number of strategies. Still, it has more strategies than direct revelation mechanism, which is the largest number of messages that is ever needed for realization.

Basically, the main idea behind the procedure of Hurwicz and Reiter (2008) is to group similar preferences behind the same equilibrium strategy. We

[26] Since $f(\Theta) = \{z, x\}$, agent 2 would be completely indifferent about any two strategies in the direct revelation mechanism.

have seen, in Example 3, that the usefulness of this result depends very heavily on the context. That is, on the exact form of the preference domain. Unfortunately, the construction does not give us any means to calculate numerical bounds for the number of strategies needed in a minimal implementing game form, and hence, there is no way to say how much it differs from the order of magnitude of the preference domain. However, the results that we have obtained stand in a stark contrast with the results obtained for other equilibrium concepts, such as Nash equilibrium. Williams (1984) and Reichelstein and Reiter (1988) have shown that the increase in communication requirements can be quite substantial in this case.

Agent 2

		s_2^1	s_2^2
	s_1^1	z	x
Agent 1	s_1^2	x	x
	s_1^3	x	y

FIGURE 5. A game form implementing f

References

Barberà, S. (2001), An Introduction to Strategy-Proof Social Choice Functions. *Social Choice and Welfare* 18, 619-653.

Bergemann, D. and Morris, S. (2005), Robust Mechanism Design. *Econometrica* 73, 1771-1813.

Dasgupta, P., Hammond, P. and Maskin, E. (1979), The Implementation of Social Choice Rules – Some General Results on Incentive Compatibility. *The Review of Economic Studies* 46, 185-216.

Dummet, M. and Farquharson, R. (1961), Stability in Voting. *Econometrica* 29, 33-44.

Gibbard, A. (1973), Manipulation of Voting Schemes: A General Result. *Econometrica* 41, 587-601.

Hall, M. (1948), Distinct Representatives of Subsets. *Bulletin American Mathematics Society* 54, 922-926.

Hurwicz, L. and Reiter, S. (2008), *Designing Economic Mechanisms*. Cambridge University Press, New York, USA.

Hurwicz, L. (1994), Economic Design, Adjustment Processes, Mechanisms, and Institutions. *Economic Design* 1, 1-14.

——— (1972), On Informationally Decentralized Systems. In *Decision and Organization*, ed. by C. B. McGuire and R. Radner. Amsterdam, North-Holland.

Maskin, E. and Moore, J. (1999), Implementation and Renegotiation. *The Review of Economic Studies* 66, 39-56.

Maskin, E. (1979), Incentive Schemes Immune to Group Manipulation. Mimeo.

Mizukami, H. and Wakayama, T. (2007), Dominant Strategy Implementation in Economic Environments. *Games and Economic Behavior* 60, 307-325.

Moulin, H. and Shenker, S. (1992), Serial Cost Sharing. *Econometrica* 60, 1009-1037.

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

——— (1980), On Strategy-Proofness and Single Peakedness. *Public*

Choice 35, 437-455.

Mount, K. and Reiter, S. (1974), The Informational Size of Message Space. *Journal of Economic Theory* 8, 161-192.

Osana, H. (1978), On the Informational Size of Message Space for Resource Allocation Processes. *Journal of Economic Theory* 17, 66-78.

Oström, E. (1986), An Agenda for the Study of Istitutions. *Public Choice* 48, 3-25.

Reichelstein, S. and Reiter, S. (1988), Game Forms with Minimal Strategy Space. *Econometrica* 56, 661-692.

Repullo, R. (1985), Implementation in Dominant Strategies under Complete and Incomplete Information. *The Review of Economic Studies* 52, 223-229.

Saijo, T., Sjöström, T. and Yamato, T. (2007), Secure Implementation. *Theoretical Economics* 2, 204-229.

Satterthwaite, M. and Sonnenschein, H. (1981), Strategy-Proof Allocation Mechanisms at Differentiable Points. *Review of Economic Studies* 48, 587-597.

Satterthwaite, M. (1975), Strategy-Proofness and Arrows' Condition: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *Journal of Economic Theory* 10, 187-217.

Segal, I. (2007), The Communication Requirements of Social Choice Rules and Supporting Budget Sets. *Journal of Economic Theory* 136, 341-378.

Williams, S. (1984), Realization and Nash Implementation – Two Aspects of Mechanism Design. *Econometrica* 54, 139-51.

ARTICLE 2

Ville Korpela (2010) Nash implementation theory – a note on full characterizations. *Economics Letters*, **108**, pp. 283-285.

Economics Letters 108 (2010) 283–285

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Nash implementation theory — A note on full characterizations

Ville Korpela *

Public Choice Research Centre (PCRC), Publicum, Assistentinkatu 7, FIN-20014, University of Turku, Finland

ARTICLE INFO

Article history:
Received 15 October 2009
Received in revised form 25 January 2010
Accepted 20 May 2010
Available online 27 May 2010

JEL classification:
C72
D71

Keywords:
(Maskin) Monotonicity
Moore–Repullo-set
Nash implementation
No veto power
$n \geq 2$ players

ABSTRACT

The first full characterization of social choice correspondences that are implementable in Nash equilibrium, given in Moore and Repullo (1990), together with the working principle behind an algorithm to check this condition, given in Sjöström (1991), can be used to give a simple necessary and sufficient condition for implementation that is a generalization of monotonicity.

1. Introduction

The first important result on Nash implementation was derived by Eric Maskin in 1977 and later published as Maskin (1999). It states that *monotonicity*, together with *no veto power*, forms a sufficient condition for a social choice correspondence to be implementable in Nash equilibrium. This simple partial characterization has one disturbing feature. It does not have bite with some very simple social choice correspondences, such as the constant or dictatorial choice correspondence. At first glance this does not seem too problematic as we can easily construct an implementing game form in both cases. But consider a choice correspondence that is monotonic, does not satisfy no veto power, and, is not Nash implementable. This time there is nothing we can do, except to demonstrate that no implementing game form can in fact exist.

This difficulty can only be avoided through a full characterization, like the one in Moore and Repullo (1990).[1] Their necessary and sufficient condition is based on the existence of certain sets that can be understood as representing the alternatives on specific rows of the implementing game form. How to construct these sets is obvious when monotonicity and no veto power are satisfied. It is also clear in the case of constant or dictatorial choice correpondence, which just reflects the fact that an implementing game form is easy to construct. However, exactly the same social choice correspondences that were problematic before, are still problematic. To verify that the desired sets do not exist is essentially analogous to verifying that no implementing game form does exist.

In comparison with the partial characterization of Maskin (1999), there is one considerable advantage in the full characterization of Moore and Repullo (1990). An algorithm to check whether a given social choice correspondence is Nash implementable can be based on the latter. This was accomplished in Sjöström (1991). The central working principle of this algorithm is to form test sets that must be accepted as Moore–Repullo-sets exactly in the case of social choice correspondence is Nash implementable. A simple necessary and sufficient condition for Nash implementation, that may be seen as a generalization of (Maskin) monotonicity, can be based on this idea.

Our main purpose is to express this monotonicity condition explicitly. The main contribution of this paper is then to compress the full characterization of Moore and Repullo (1990), which consists of three separate conditions, into a one unified condition that should be a lot easier to comprehend.

The organization of the paper is as follows. In Section 2 we fix notation and run through the main concepts. Section 3 elaborates on the new necessary and sufficient condition for a social choice correspondence to be Nash implementable. This will be termed as *generalized monotonicity*. Also Danilov (1992) has presented a stronger monotonicity condition, but this is necessary and sufficient only in comprehensive domains.[2] We shall make no such restrictions. Section 4 concludes.

* Tel.: +358 2 333 6284; fax: +358 2 333 5893.
E-mail address: ville.korpela@utu.fi.

[1] Dutta and Sen (1991) have independently derived the same condition in two player case.

[2] Yamato (1992) generalized this result, but also he made a restriction on the class of admissible domains called *condition D*.

0165-1765/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2010.05.019

2. Preliminaries

Let $N=\{1,\ldots n\}$ be the set of players and A the non-empty set of social alternatives. Let $\mathscr{R}_A$ be the class of all complete and transitive preference orderings on A *i.e.* the unrestricted domain. *A social choice correspondence* (SCC) f is defined to be any function $f:\mathscr{R}\to 2^A\backslash\{\emptyset\}$, such that $\mathscr{R}\subseteq\times\mathscr{R}_A$. A SCC gives the set of socially acceptable alternatives as a function of the prevailing preference profile. We shall denote a generic element of $\mathscr{R}$ by $\boldsymbol{R}$ and the preference of player i in this profile by $\boldsymbol{R}_i$. In the usual manner, $\boldsymbol{R}_{-i}$ will denote an $(n-1)$-dimensional vector that specifies a preference ordering for all players except i.

Two properties of SCCs' are needed. Let $L_i(a,\ \boldsymbol{R})$ denote the *lower contour set* of player i at $a\in A$ when the preference profile is $\boldsymbol{R}$ *i.e.* $L_i(a,\boldsymbol{R})=\{b\in A|a\boldsymbol{R}_ib\}$.[3] A SCC f is called *monotone*, if $a\in f(\boldsymbol{R})$ and $L_i(a,\ \boldsymbol{R})\subseteq L_i(a,\ \boldsymbol{R}')$ for all $i\in N$ implies $a\in f(\boldsymbol{R}')$. A SCC f satisfies *no veto power*, if $\#\{i\mid L_i(a,\ \boldsymbol{R})=A\}=n-1$ implies $a\in f(\boldsymbol{R})$ for all $a\in A$. Moreover, the set of maximal elements in $C\subseteq A$ with respect to $\boldsymbol{R}_i$ is denoted by $M_i(C,\boldsymbol{R})$ *i.e.* $M_i(C,\boldsymbol{R})=\{a\in C|a\boldsymbol{R}_ic$ for all $c\in C\}$.

A tuple $G=(S,g)$, where $S=\times S_i$ is the *set of strategy profiles* and g: $S\to A$ is an *outcome function*, is called a *game form* or a *mechanism*. Here S_i is the strategy set of player i. When a preference profile $\boldsymbol{R}$ is given, this game form will define a game $\Gamma=\langle G,\boldsymbol{R}\rangle$ in normal form. A Nash equilibrium of this game is a strategy profile $s\in S$ such that $g(s)$ $\boldsymbol{R}_ig(s_i^*,s_{-i})$ for all $s_i^*\in S_i$ and all $i\in N$. Let $NE(G,\boldsymbol{R})$ be the set of all Nash equilibria in the game Γ. We say that the game form *G Nash implements* a SCC f, if $g(NE(G,\boldsymbol{R}))=f(\boldsymbol{R})$ for all $R\in\mathscr{R}$. If there exists a game form G that Nash implement f, then f is called *Nash implementable*.

3. Generalization of monotonicity — a necessary and sufficient condition

The test sets that are used in the algorithm of Sjöström (1991) can be formed iteratively. To this end, let's define an operator[4]

$$\rho(A)=A\backslash\{a\in A\backslash f(\mathscr{R})\,|\,\exists\mathscr{R}\in\boldsymbol{R}:a\in M_i(A,\boldsymbol{R})\forall i\in N\}$$

that excludes some of the alternatives which can never be used in the implementing game form.[5] The set of alternatives that can *possibly* be used is then defined iteratively as $A^*=\cap\rho^m(A)$, since a one round elimination may create new alternatives that can not be used.[6,7] From now on $L_i^*(a,\boldsymbol{R})$ will denote the set $L_i(a,\boldsymbol{R})\cap A^*$.

We shall also need another operator

$$\delta_i(C)=C\backslash\left\{a\,|\,\exists\boldsymbol{R}\in\mathscr{R}:a\notin f(\boldsymbol{R}),a\in M_i(C,\boldsymbol{R}),a\in M_j(A^*,\boldsymbol{R})\forall j\in N\backslash i\right\},$$

which for a given set $C\in A$ and for a given player $i\in N$, eliminate all veto problematic alternatives.[8] Iterating δ_is, in the same fashion we did with ρ, allow us to define new sets in the following way $\hat{L}_i^*(a,\boldsymbol{R})=\cap\delta_i^m(L_i^*(a,\boldsymbol{R}))$.[9]

Definition. A SCC f satisfy *generalized monotonicity*, if for all $\boldsymbol{R},\boldsymbol{R}'\in\mathscr{R}$ and all $a\in f(\boldsymbol{R})$, we have $a\in\hat{L}_i^*(a,\boldsymbol{R})$, and the following condition holds

$$\hat{L}_i^*(a,\boldsymbol{R})\subseteq L_i^*(a,\boldsymbol{R}')\forall i\in N\quad\Rightarrow\quad a\in f(\boldsymbol{R}')^{10}$$

The proof that generalized monotonicity is both a necessary and sufficient condition for a SCC to be implementable in Nash equilibrium when there are three or more players is simple. It follows directly from the characterization of Moore and Repullo (1990), called *condition* μ, together with the working principle behind the algorithm in Sjöström (1991).

Condition μ. There exists a set $B\subseteq A$, and for each $j\in N$, $\boldsymbol{R}\in\mathscr{R}$ and $a\in f(\boldsymbol{R})$, there exists a set $C_i(a,\boldsymbol{R})\subseteq B$ with $a\in M_i(C_i(a,\boldsymbol{R}),\boldsymbol{R})$, such that for all $\boldsymbol{R}'\in\mathscr{R}$ the following conditions are satisfied:

(*i*) If $a\in\cap M_i(C_i(a,\boldsymbol{R}),\boldsymbol{R}')$, then $a\in f(\boldsymbol{R}')$.
(*ii*) If $c\in M_i(C_i(a,\boldsymbol{R}),\boldsymbol{R}')\cap[\cap_{j\neq i}M_j(B,\boldsymbol{R}')]$ for some $i\in N$, then $c\in f(\boldsymbol{R}')$.
(*iii*) If $d\in\cap M_i(B,\boldsymbol{R}')$, then $d\in f(\boldsymbol{R}')$.

Theorem. *Let $n\geq 3$. A SCC f is Nash implementable if and only if it satisfy generalized monotonicity.*

Proof. The algorithm in Sjöström (1991) is based on the fact that a SCC f is Nash implementable exactly in the case that the choice $B=A^*$ and $C_i(a,\ \boldsymbol{R})=\hat{L}_i^*(a,\ \boldsymbol{R})$ for all $i\in N$ and all $\boldsymbol{R}\in\mathscr{R}$, such that $a\in f(\boldsymbol{R})$, satisfies condition μ. But with these sets condition (i) is generalized monotonicity, so f is in fact Nash implementable exactly in the case that generalized monotonicity holds (as conditions (ii) and (iii) are satisfied by construction). □

Example 1. Let $c\in A$ and define the SCC $f:\mathscr{R}_A\to A$ by $f(\boldsymbol{R})=\{a\in A|a\boldsymbol{R}_ic\forall i\in N\}$. This correspondence is monotone, but does not satisfy no veto power. Now $A^*=A$ and $\hat{L}_i^*(a,\ \boldsymbol{R})=\emptyset$, so f is not Nash implementable by generalized monotonicity.

Example 2. Let $c\in A$ and assume that $f_c:\mathscr{R}\to A$ is the constant SCC *i.e.* $f(\mathscr{R})=\{c\}$. In this case A^* and $\hat{L}_i^*(a,\boldsymbol{R})$ may be hard to find, since they depend heavily on the preference domain $\mathscr{R}$. Still, generalized monotonicity is satisfied trivially — the right hand side of the implication is always true.

Generalized monotonicity is a necessary condition also in the two player case. This can be seen in the following way. First, if B and the class $\{C_i(a,\boldsymbol{R})\}$ satisfy condition μ, then it must definitely be that $B\subseteq A^*$ and $C_i(a,\ \boldsymbol{R})\subseteq\hat{L}_i^*(a,\ \boldsymbol{R})$ for all $i\in N$ and all $\boldsymbol{R}\in\mathscr{R}$, such that $a\in f(\boldsymbol{R})$. In other words, A^* and $\{\hat{L}_i^*(a,\boldsymbol{R})\}$ form the class of maximal sets that can possibly satisfy condition μ. Second, if these sets do not satisfy condition μ, then no subsets of them can either.[11] But the implementation condition for the two player case, condition $\mu 2$ in Moore and Repullo (1990), require that also condition μ is satisfied. Hence generalized monotonicity is necessary also in the two player case.

4. Conclusion

Many characterizations of Nash implementable social choice correspondences have been given in the literature. Simple partial characterizations usually exclude most of the hard and/or easy cases, and on the other hand, full characterizations tend to be complicated because they try to incorporate everything. The latter also goes for generalized monotonicity, as it cannot always be easily verified. One

---

[3] Strictly speaking L_i only depends on $\boldsymbol{R}_i$. Still, it is notationally more convenient to express it as depending on $\boldsymbol{R}$.

[4] For simplicity we do not express the operator as explicitly depending on the SCC f.

[5] If an alternative in $A\backslash f(\mathscr{R})$ is top ranked for every player, it would clearly constitute a bad Nash equilibrium in any game form.

[6] $\rho^m(A)$ is defined recursively as: $\rho^0(A)=A$ and $\rho^h(A)=\rho(\rho^{h-1}(A))$ for all $h\geq 1$.

[7] With most SCCs' there is no difference between A and A^*. The set A^* is just defined to obtain full generality. In fact, these sets can differ only when unanimity does not hold.

[8] Veto problematic in the sense that elements in a row of an implementing game form can not coincide with the set C, since there would be a bad Nash equilibrium.

[9] Notice that we may not be able to construct the set $\hat{L}_i^*(a,\boldsymbol{R})$ in algorithmic sense, but it does exist. Same goes with A^*.

[10] If no veto power is satisfied, then $\hat{L}_i^*(a,\boldsymbol{R})=L_i^*(a,\boldsymbol{R})=L_i(a,\boldsymbol{R})$ and the definition of generalized monotonicity coincide with monotonicity.

[11] Since these sets satisfy (ii) and (iii) of condition μ by construction, the failure must come from (i). But if this is the case, then one can easily see that no subsets of them can satisfy (i) either.

possible way to get around this would be to give the characterizing condition in multiple parts, each part characterizing Nash implementable SCCs in different types of preference domains. By this we do not mean a simple restriction on the admissible domain, rather, we mean a carful division of the class of all domains in parts that are similar from Nash implementation point of view. Still, it seems that generalizing monotonicity is the right way to go. Some papers, like Benoît and Ok (2006), have chosen to generalize no veto power. But as we have shown, the main purpose of this property is just to simplify certain iterative processes.

Acknowledgement

Many thanks to Hannu Vartiainen for useful discussions.

References

Benoît, J.-P., Ok, E.A., 2006. Maskin's theorem with limited veto power. Games and Economic Behavior 55, 331–339.
Danilov, V., 1992. Implementation via Nash equilibrium. Econometrica 60, 43–56.
Dutta, B., Sen, A., 1991. A necessary and sufficient condition for two-person Nash implementation. Review of Economic Studies 58, 121–128.
Maskin, E., 1999. Nash equilibrium and welfare optimality. Review of Economic Studies 66, 23–38.
Moore, J., Repullo, R., 1990. Nash implementation — A full characterization. Econometrica 58, 1083–1099.
Sjöström, T., 1991. On the necessary and sufficient conditions for Nash implementation. Social Choice and Welfare 8, 333–340.
Yamato, T., 1992. On Nash implementation of social choice correspondences. Games and Economic Behavior 4, 484–492.

ARTICLE 3

Ville Korpela (2011) Implementation without rationality assumptions. Forthcoming in *Theory and Decision* at the time of press. Published online: DOI 10.1007/s11238-011-9251-x.

Theory Dec.
DOI 10.1007/s11238-011-9251-x

Implementation without rationality assumptions

Ville Korpela

Abstract Hurwicz (Social Choice and public decision making. Essays in honor of Kenneth J. Arrow, Cambridge University Press, Cambridge, 1986) was the first to study an approach to implementation theory based on choice functions instead of preference relations. We argue that the solution concept used by him, the generalized Nash equilibrium, is not really compatible with the idea that individual behavior is describable by a choice function. A new solution concept that better fits the choice function framework is then introduced. Using this, we investigate what behavioral assumptions are needed for the full characterizations of Nash implementable social choice correspondences to still hold. We will show that a condition known as Property α is central.

Keywords Moore–Repullo set · Nash Implementation · $n \geq 2$ Players · Properties α and γ

JEL Classification C72 · D03 · D71 · D78

1 Introduction

Complete and transitive preferences are assumed in the implementation literature, with very few exceptions. This is despite the fact that there is nothing in the idea of mechanism design that really requires that assumption. After all, the goal is to design an institutional setting that produces socially acceptable outcomes as a function of plausible behavior. The main message of Hurwicz (1986) is that the important results

V. Korpela (✉)
Public Choice Research Centre (PCRC), University of Turku, Publicum,
Assistentinkatu 7, 20014 Turku, Finland
e-mail: ville.korpela@utu.fi

Published online: 27 March 2011

of Maskin (1977)[1] on Nash implementation are still essentially valid, even if we do not make the assumption of complete and transitive preferences.

In conventional implementation theory, the domain of a social choice correspondence is a subset of all possible profiles of preference relations. The central idea of Hurwicz (1986) was to replace this domain with a subset of all possible profiles of choice functions. To do this, a new solution concept had to be defined. The one used by Hurwicz (1986) is entitled generalized Nash equilibrium (GNE).

However, GNE is not really compatible with the idea that individual behavior is describable by a choice function. More specifically, GNE is based on a pairwise comparison of alternatives in the rows of a game form, whereas a truly choice function-based solution concept would select from the set of all alternatives in a row simultaneously. For the former to imply the latter, an additional assumption, called *normality* in Sen (1977), has to be made. This condition states that choice function must be generated by the "base relation", also known as the "underlying binary relation" (Mas-Colell et al. 1995), that it defines. Unfortunately, this restrictive assumption precludes a large class of behaviors.

The main purpose of this article is to generalize the result of Hurwicz (1986). We define a new solution concept, which we dub behavioral Nash equilibrium (BehNE). Contrary to Hurwicz (1986), this does not place any restrictions on the choice behavior. We prove that the new solution concept, together with Sen's Property α, can be used to fully characterize implementable social choice correspondences directly in terms of choice functions. Property α says, roughly, that whenever an alternative is selected in one situation, it must also be selected in any situation that has less alternatives. Since no additional assumptions are made concerning the behavior of players, our result is a generalization of the characterization by Moore and Repullo (1990).

Furthermore, it is easy to express the exact scope of the generalization. This is due to the fact that normality is not equivalent to Property α alone. Another property, known as Property γ, is needed (Sen 1977, Proposition 8). Property γ says, roughly, that whenever an alternative is selected in multiple different situations, it must also be selected in a situation where all the alternatives are possible at once.

Our study contributes to the recent trend in implementation theory which incorporates more flexible behavioral assumptions into the theory (Matsushima 2008a,b). At the same time, it has become increasingly apparent that *framing*, or *menu-dependence*, is an essential feature of human behavior (Rubinstein and Salant 2008; Manzini and Mariotti 2007). This is exactly what Properties α and γ are about; they both rule out certain types of framing. In effect, we will be able to handle certain forms of framing (γ need not hold), but not all (α must hold). To dispense with the latter form of framing is beyond the scope of this article, even though an important topic for future research. In a real world, framing usually happens in a situation where people behave in what could be called a *rule of thumb* manner. We will give an example of such behavior in Sect. 4.

Independently of this article, Ray (2010) has recently studied social choice correspondences that are implementable without assuming full rationality. To the best of our

[1] This working paper was later published as Maskin (1999).

knowledge, this is the only other paper on the subject in addition to Hurwicz (1986). Ray (2010), however, is mostly interested in generalizing the necessary and sufficient conditions of Maskin (1977), whereas we are interested in the full characterization due to Moore and Repullo (1990).

It should be noted that by the very definition of implementation, questions about the existence of equilibria are irrelevant, even when games are played via choice functions. Implementation can only be successful if a game form *does* exist that has at least one equilibrium for all possible profiles of choice function.

Finally, we note that Ray (2010) also provides a generalization of the Moore–Repullo characterization. However, in contrast to this paper, his aim is to simultaneously solve the *equilibrium selection problem*. The definition of implementation allows a game form to have multiple equilibria for the same preference profile, which can cause a coordination failure. If we require that this can never happen, then the set of implementable social choice correspondences can, at least in principle, be reduced. Interestingly, this problem has not been solved satisfactorily, even in the fully rational case. Our approach is to focus on Sjöström (1991) algorithm to check the Moore–Repullo characterization. We will show that this algorithm still works if normality is assumed, but may otherwise become more inefficient. To conclude, this article and Ray (2010) have independently tackled the same problem from slightly different points of view.

The article is organized as follows: Section 2 introduces the basic notation and concepts. Section 3 argues for the new solution concept that does not necessitate any additional assumptions on choice behavior in contrast to Hurwicz (1986). Section 4 provides the characterization result and explains why Property α is needed. Before Sect. 6 concludes this article, Sect. 5 investigates whether the verification of the characterizing condition is as simple as in the case of complete and transitive preferences.

2 Notation and definitions

Let $I = \{1, \ldots, n\}$ be the set of players and X the set of social alternatives. A typical element of I is denoted by i or j, and a typical element of X by x, y or z, and so forth. A *choice function* is any function $C\colon 2^X \setminus \{\emptyset\} \to 2^X$, such that $C(A) \subseteq A$ for all $A \in 2^X \setminus \{\emptyset\}$. Choice function may equally well be an individual decision function or a group decision function. We do not preclude the case $C(A) = \emptyset$. This can be interpreted by saying that *deadlocks*, where a player refuses to make any choice at all, are allowed. A binary relation $\succeq$, defined by the rule[2]

$$x \succeq y \text{ if and only if } x \in C(x, y),$$

is called a *base relation* of the choice function C. This relation may not be complete. If $\succeq$ is able to generate the choice behavior C, that is, if $C(A) = \{x \in A \mid x \succeq y \text{ for all } y \in A\}$ for all $A \in 2^X \setminus \{\emptyset\}$, then we say that choice function C is *normal* (Sen 1977, Proposition 7).

[2] We will use the standard convention $C(\{x, y\}) = C(x, y)$ throughout.

The set of all possible choice functions over X is denoted by $\mathscr{C}_X$. Any function $f: \mathscr{C} \to 2^X \setminus \{\emptyset\}$, such that $\mathscr{C} \subseteq \prod_{i \in I} \mathscr{C}_X$, is called a social choice correspondence (SCC). Since no assumptions about rationality have been made, we interpret this by saying that social choice is based on behavioral patterns. A typical element of $\mathscr{C}$ is denoted by $\boldsymbol{C}$ and the choice function of player i in this profile by $\boldsymbol{C}_i$. In addition, $\boldsymbol{C}_{-i}$ denotes the $(n-1)$-dimensional profile that specifies a choice function for all players except i and $\mathscr{C}_i = \{C \in \mathscr{C}_X \mid C = \boldsymbol{C}_i \text{ for some } \boldsymbol{C} \in \mathscr{C}\}$.

A tuple $G = (S, g)$, where $S = S_1 \times \cdots \times S_n$ is the *set of strategy profiles* and $g: S \to X$ is the *outcome function*, is called a *game form* or a *mechanism*. For a fixed profile of choice functions $\boldsymbol{C} = (C_1, \ldots, C_n)$, this game form will define a game $\Gamma(\boldsymbol{C}) = (G; \boldsymbol{C})$. The specification of choice functions is enough to fully describe the behavior of all players when facing any row of the game form. Let E be some game theoretic solution concept. The set $E[\Gamma(\boldsymbol{C})] \subseteq S$ consists of all strategy profiles that are solutions in the game $\Gamma(\boldsymbol{C})$ with respect to E. The game form G is said to *implement f in E* if

$$g(E[\Gamma(\boldsymbol{C})]) = f(\boldsymbol{C}) \ \text{ for all } \ \boldsymbol{C} \in \mathscr{C}.$$

If some game form G implements f in E, then we call this SCC *implementable in E*.

3 A new solution concept for the implementation approach based on choice functions

There are several ways to generalize the idea of Nash equilibrium for the game $\Gamma(\boldsymbol{C})$. The first one is from Hurwicz (1986)

Definition 1 Strategy profile $\boldsymbol{s}^* \in S$ is a generalized Nash equilibrium (GNE) of $\Gamma(\boldsymbol{C})$ if and only if

$$g(\boldsymbol{s}^*) \in \boldsymbol{C}_i(g(\boldsymbol{s}^*), g(\boldsymbol{s}^*_{-i}, s_i)) \text{ for all } i \in I \text{ and all } s_i \in S_i. \tag{1}$$

In this definition, players are comparing alternatives in the rows of a game in a pairwise manner. The definition is not really compatible with the idea of a choice function unless we assume normality.[3] To understand this, consider the set $g(\boldsymbol{s}^*_{-i}, S_i) = \{g(\boldsymbol{s}^*_{-i}, s_i) \mid s_i \in S_i\}$. This set includes all alternatives that player i can obtain by unilaterally deviating from the GNE $\boldsymbol{s}^*$. If condition (1) truly determines an equilibrium of the game $\Gamma(\boldsymbol{C})$, then it must guarantee that $g(\boldsymbol{s}^*) \in \boldsymbol{C}_i(g(\boldsymbol{s}^*_{-i}, S_i))$ for all $i \in I$. In other words, the pairwise comparisons made in condition (1) must be sufficient to guarantee that $g(\boldsymbol{s}^*)$ is chosen when considering the whole set $g(\boldsymbol{s}^*_{-i}, S_i)$ at once. This is why

[3] Hurwicz (1986) did know this. His main goal was to handle choice behavior that can have *Condorcet cycles*, and this is what he tried to take into account when defining GNE. Normal choice functions can have Condorcet cycles. However, all choice functions that have Condorcet cycles are not normal. If $C(x, y) = x$, $C(y, z) = y$, $C(x, z) = z$ and C is normal, then we must have $C(\{x, y, z\}) = \emptyset$. That is, there must be a deadlock when choosing from the Condorcet cycle. The assumption of normality is, of course, a generalization in comparison to the standard case, where the *Weak Axiom of Revealed Preference* (WARP) is assumed. See Moulin (1985), Plott (1976), Samuelson (1938), and Sen (1970) for more on WARP.

the consistency of condition (1) necessitates an additional assumption of normality. In addition to this implicit restriction, Definition 1 has another shortcoming. That is, even though $g(\boldsymbol{s}^*)$ does not win every alternative of $g(\boldsymbol{s}^*_{-i}, S_i)$ in a pairwise comparison, it may very well be chosen from the set $g(\boldsymbol{s}^*_{-i}, S_i)$. The following two examples will clarify this.

Example 1 [4]Let $X = \{x, y, z\}$ and $\mathscr{C} = \{C_1, C_1'\} \times \{C_2\}$, so that player 1 has two possible choice functions and player 2 only one possible choice function. Assume that $C(x, y) = \{x\}, C(y, z) = \{y\}$ and $C(z, x) = \{z\}$ for both $C \in \{C_1, C_1'\}$. This implies that the base relation $\succeq$ is cyclic for both C_1 and C_1', that is, $x \succeq y \succeq z \succeq x$. To complete the description, let $C_1(X) = x$ and $C_1'(X) = y$. The exact form of C_2 will play no role. Now, consider an SCC f defined by the rule $f(C_1, C_2) = x$ and $f(C_1', C_2) = y$. This SCC should be implementable − the designer can simply let player 1 select from the set X. Despite this, a game form that can implement f in GNE does not exist. Assuming the contrary, a game form $G = (S, g)$ implements f in GNE. This implies that a GNE $\boldsymbol{s}^*$ of $\Gamma(C_1, C_2)$ must exist, such that $g(\boldsymbol{s}^*) = x$. Moreover, it must be that $z \notin g(S_1, \boldsymbol{s}^*_2)$, otherwise $C_1(x, z) = z$ would contradict (1). This implies that $g(S_1, \boldsymbol{s}^*_2) \subseteq \{x, y\}$ and hence that $\boldsymbol{s}^*$ is also a GNE in the game $\Gamma(C_1', C_2)$. This is impossible, since $f(C_1', C_2) = y$. Therefore, a game form that implements f in GNE cannot exist. This is despite the fact that f should be deemed implementable. □

This example is not unrealistic. It has been pointed out many times that even rational agents can have cyclic preferences, at least when alternatives are evaluated in multiple dimensions (Bar-Hillel and Margalit 1988; Manzini and Mariotti 2007; Tversky 1969). In the next example, we show that the result of Example 1 does not crucially depend on cyclicity.

Example 2 The environment is exactly as in Example 1, except that this time $C(x, y) = \{x, y\}, C(x, z) = \{x\}$ and $C(y, z) = \{y\}$ for both $C \in \{C_1, C_1'\}$. We interpret this as follows: Player 1 is indecisive between alternatives x and y, but the presence of alternative z would make the player choose between them. For exactly the same reason as before, SCC f should be implementable. Again, it is not implementable in GNE. Assume that a game form $G = (S, g)$ implements f. Let $\boldsymbol{s}^*$ be a GNE of the game $\Gamma(C_1, C_2)$, such that $g(\boldsymbol{s}^*) = x$. It is obvious, by condition (1), that $\boldsymbol{s}^*$ must be a GNE also in the game $\Gamma(C_1', C_2)$. This cannot be. Hence, a game form that implements f in GNE does not exist. □

There is a subtle difference between Examples 1 and 2 that will be explained in due course. For the time being, notice that both examples have choice functions that are not normal. These examples work as a motivation for the second generalization[5] of Nash equilibrium.

[4] The fact that there are only two players is not critical. We could simply add "dummy" players identical to player 2.

[5] Ray (2010) calls it a setwise Nash equilibrium.

Definition 2 Strategy profile $s^* \in S$ is a behavioral Nash equilibrium of $\Gamma(\boldsymbol{C})$ if and only if

$$g(s^*) \in \boldsymbol{C}_i(g(s^*_{-i}, S_i)) \text{ for all } i \in I. \tag{2}$$

In this definition, instead of making pairwise comparisons, players are now choosing simultaneously from all the alternatives in the rows of a game. It has been known since Hurwicz (1986) that the important results of Maskin (1977) on Nash implementation are still valid if we use GNE as the solution concept. Naturally, this requires, in addition to the normality of choice functions as we have argued, that the conditions known as *Maskin-monotonicity* and *No Veto Power* are generalized to a choice function setting.[6] The main contribution of Hurwicz (1986) was to demonstrate that central implementation results do not essentially depend on the assumption of complete and transitive preferences, or in other words, on the rationality of behavior. This makes perfect sense. There is nothing in the mechanism design approach that would crucially depend on rationality per se. After all, the purpose is simply to design an institutional setting that always leads to desired outcomes as a function of plausible behavior. Based on this intuition, we are going to show that it is possible to characterize the set of SCCs that are implementable in BehNE, and hence broaden the scope of implementation theory even further.

4 A characterization of social choice correspondences implementable in behavioral Nash equilibrium

The full characterization of Nash implementable SCCs, given in Moore and Repullo (1990),[7] is based on the existence of a certain class of sets. These sets are, in fact, rows of a possible game form. The characterization places restrictions on the best alternatives in these sets. These are the alternatives, that a rational player would choose from the row. This raises a natural question: If we impose the same restrictions on any possible behavior, rational or not, does the result still hold true? This motivates us to generalize *Condition* μ of Moore and Repullo (1990) as follows:

Condition (λ) . There is a set $Y \subseteq X$, and for each $i \in I$, $\boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$, there is a set $R_i(x, \boldsymbol{C}) \subseteq Y$ with $x \in \boldsymbol{C}_i(R_i(x, \boldsymbol{C}))$, such that for all $\boldsymbol{C}^* \in \mathscr{C}$, the following three conditions are satisfied:

(i) If $x \in \bigcap \boldsymbol{C}^*_i(R_i(x, \boldsymbol{C}))$, then $x \in f(\boldsymbol{C}^*)$.

(ii) If $y \in \boldsymbol{C}^*_i(R_i(x, \boldsymbol{C})) \bigcap \left[\bigcap_{j \neq i} \boldsymbol{C}^*_j(Y)\right]$ for some $i \in I$, then $y \in f(\boldsymbol{C}^*)$.

(iii) If $z \in \bigcap_{i \in I} \boldsymbol{C}^*_i(Y)$, then $z \in f(\boldsymbol{C}^*)$.

[6] SCC f satisfies *Maskin-monotonicity*, if for all $\boldsymbol{C}, \boldsymbol{C}' \in \mathscr{C}$, such that $x \in f(\boldsymbol{C})$, the following holds: If $x \in \boldsymbol{C}_i(x, y)$ implies $x \in \boldsymbol{C}'_i(x, y)$ for all $i \in I$ and all $y \in X$, then $x \in f(\boldsymbol{C}')$. It satisfies *No Veto Power*, if the following holds true for all subset I^* of I that have either n or $n-1$ players: If $y \in C_i(y, z)$ for all $i \in I^*$ and all $z \in X$, then $y \in f(\boldsymbol{C})$. Remember that these properties together form a sufficient condition for implementation (now in GNE), not a necessary one. How these properties should be generalized when normality *is not* assumed can be found in Ray (2010).

[7] This characterization was derived independently by Dutta and Sen (1991) in the two-player case.

Items (i),(ii) and (iii) can be interpreted as follows: (i) *Maskin-monotonicity*, (ii) *No Veto Power* and (iii) *Unanimity*. These properties have simply been generalized by demanding that they hold in a suitable class of sets. The sufficiency of Condition λ can be proved without *any* restrictions on the choice behavior in exactly the same way as the sufficiency of Condition μ in Moore and Repullo (1990). It is the necessity side that does not hold universally. In the original characterization of Moore and Repullo (1990), the existence of the set Y is crucially dependent on the fact that rational behavior allows for some degree of predictability. In their necessity proof, if $g: S \to X$ implements f, then Y can be chosen as $g(S)$ to satisfy (iii). Assume that x is the best alternative in $g(S)$ for all players under some preference profile. Then, the strategy profile $\boldsymbol{s} \in S$, such that $g(\boldsymbol{s}) = x$, must be an equilibrium of g. This implies that x must belong to the range of f under this preference profile as required by (iii). For this reason, we have to assume something about the choice functions in $\bigcup_{i\in I} \mathscr{C}_i$ so that this still holds true. The next property, which is from Sen (1977), turns out to be central.[8]

Property α For all $A, B \subseteq X$ and all $x \in X$, if $x \in A \subseteq B$ and $x \in C(B)$, then it is also true that $x \in C(A)$.

Property α is concerned with keeping a chosen alternative x still choosable once the set has been shrunk by dropping other alternatives. In other words, certain types of framing are not allowed. The assumption is not completely harmless. Kahneman and Tversky (1984), among others, have found that framing does happen in practice, and by assuming Property α we will exclude some of these cases. Nonetheless, we have the following two theorems.

Theorem 1 (Sufficiency) *Let $n \geq 3$. If SCC f satisfies Condition λ, then it is implementable in BehNE.*

Proof A direct generalization of the proof in Moore and Repullo (1990). The proof is given in the Appendix. □

Theorem 2 (Necessity) *Let $n \geq 3$ and assume that all choice functions in $\bigcup_{i\in I} \mathscr{C}_i$ satisfy Property α. If SCC f is implementable in BehNE, then it satisfies Condition λ.*

Proof A simple generalization of the proof in Moore and Repullo (1990) with the help of Property α. The proof is given in the appendix. □

The reason we need Property α in the necessity part is very simple. Suppose that f is implementable by the game form $G = (S, g)$. For any $\boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$, let $\boldsymbol{s}^*(x, \boldsymbol{C})$ be some BehNE of this game form, such that $g(\boldsymbol{s}^*(x, \boldsymbol{C})) = x$. Choose Y as the set $g(S)$ of all alternatives used in the implementing game form and $R_i(x, \boldsymbol{C})$ as the set $g(S_i, \boldsymbol{s}^*_{-i}(x, \boldsymbol{C}))$ of all alternatives that player i can obtain by unilaterally deviating from the BehNE $\boldsymbol{s}^*(x, \boldsymbol{C})$. These sets must satisfy Condition λ when all choice functions in $\bigcup_{i\in I} \mathscr{C}_i$ satisfy Property α. For example, if $x \in \bigcap_{i\in I} \boldsymbol{C}_i(Y)$ for

[8] In the literature, this condition is also known as the *Chernoff property*, after Chernoff (1954), and *independence of irrelevant alternatives*, used in Nash (1950), for example.

some $\boldsymbol{C} \in \mathscr{C}$, then any strategy profile $\boldsymbol{s} \in S$, such that $g(\boldsymbol{s}) = x$, must be a BehNE of the game $\Gamma(\boldsymbol{C})$ by Property α. Hence, $x \in f(\boldsymbol{C})$ must hold as required by (iii). A similar argument applies in the case of (i) and (ii). A more detailed account is shown in the appendix. Besides the obvious fact that Hurwicz (1986) gave only a partial characterization, it is possible to express exactly the extent of the generalization that we have made using well-known properties from the literature. Normality, which is the implicit consistency requirement behind Definition 1 used by Hurwicz (1986), is more demanding than Property α alone. Normality implies that Property α must hold, but to guarantee that normality holds, we also need Property γ (Sen 1977).

Property γ For any $M \subseteq 2^X$ and $x \in X$, if $x \in C(A)$ for all $A \in M$, then it is also true that $x \in C(\cup M)$.

This property, then, identifies exactly the class of choice functions that Hurwicz (1986) was unable to deal with. If Property γ does not hold for all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$, then some choice functions in the domain are not normal, and hence it is *not* legitimate to use GNE as the solution concept. But if all choice functions in the domain satisfy Property α, even though not Property γ, we can still use BehNE as the solution concept and apply Theorem 2.

Property γ requires that alternative x, which is selected from smaller sets, must also be selected from a larger set that makes all the alternatives in these sets possible at once. Also this behavioral assumption excludes certain types of framing. The following is an archetypal example of behavior, in an economic environment, that satisfies Property α, although not Property γ.

Example 3 (A Rule of Thumb Behavior). Let $X = \{x_1, \ldots, x_m\}$ be the set of consumption goods and let $p_1, \ldots, p_m$ be the associated prices. Assume $p_1 > p_2 > \cdots > p_m$ and define $\text{Ind}(A) = \{k \mid x_k \in A\}$ for all $A \subseteq X$. Denote $k^*(A) = \max\{\text{Ind}(A)\}$ and use this to construct a choice function C by the rule

$$C(A) = \begin{cases} \left\{x_{k^*(A)},\ x_{k^*(A \setminus \{x_{k^*(A)}\})}\right\}, & \text{if } \mid A \mid > 2, \\ A, & \text{if } \mid A \mid \leq 2. \end{cases}$$

This choice function will select, from any set $A \subseteq X$, those two alternatives with the highest index. In other words, the consumer will always select the two cheapest goods when there are at least two goods available. If $x \in C(A)$, so that x has the highest or the second highest index in A, and $x \in B \subseteq A$, then x must obviously also have the highest or second highest index in B, and hence $x \in C(B)$. This means that C satisfies Property α. However, it does not always satisfy Property γ. Assume that $m \geq 3$. Let $A = \{x_{m-2}, x_m\}$ and $B = \{x_{m-2}, x_{m-1}\}$. By definition, $x_{m-2} \in C(A)$ and $x_{m-2} \in C(B)$, while $x_{m-2} \notin C(A \cup B) = \{x_{m-1}, x_m\}$. Therefore, C does not satisfy Property γ. □

Condition $\mu 2$ in Moore and Repullo (1990) can also be generalized to choice function setting as follows:

Condition $\lambda 2$ Condition λ is satisfied. Moreover, for each quadruple $(x, \boldsymbol{C}, y, \boldsymbol{C}') \in X \times \mathscr{C} \times X \times \mathscr{C}$, with $x \in f(\boldsymbol{C})$ and $y \in f(\boldsymbol{C}')$, there is an alternative $e =$

$e(x, \boldsymbol{C}, y, \boldsymbol{C}') \in R_1(x, \boldsymbol{C}) \bigcap R_2(y, \boldsymbol{C}')$, such that for all $\boldsymbol{C}^* \in \mathscr{C}$, the following condition is satisfied:

(iv) If $e \in \boldsymbol{C}_1^*(R_1(x, \boldsymbol{C})) \bigcap \boldsymbol{C}_2^*(R_2(y, \boldsymbol{C}'))$, then $e \in f(\boldsymbol{C}^*)$.

As in the case $n \geq 3$, we have the following two theorems:

Theorem 3 (Sufficiency) *Let $n = 2$. If SCC f satisfies Condition $\lambda 2$, then it is implementable in BehNE.*

Theorem 4 (Necessity) *Let $n = 2$ and assume that all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$ satisfy Property α. If SCC f is implementable in BehNE, then it satisfies Condition $\lambda 2$.*

Proof Omitted as a simple generalizations of the proofs in Moore and Repullo (1990). □

Whether these theorems are able to deal with the SCC in Examples 1 or 2 depends crucially on the specification of the domain $\mathscr{C}$. Assume that the choice function of player 2 in these examples is normal. For the SCC in Example 2, Theorem 4 implies that there must be a class of sets that satisfies Condition $\lambda 2$ and by Theorem 3, this is sufficient for implementability. For the SCC in Example 1, on the other hand, we can never apply Theorem 4. The choice function of player 1 does not satisfy Property α. Despite this, we *may* sometimes be able to verify implementability by using Theorem 3. This will ultimately depend on the choice function of player 2.

Nothing that we have presented so far is sufficient to prove that the characterization of Moore and Repullo (1990) is no longer valid when Property α ceases to hold. This property was simply used to guarantee that we can still continue to use the same *canonical mechanism* as in the original work of Moore and Repullo (1990). But another kind of canonical mechanism could, at least in principle, exist. The following example verifies that once Property α does not hold, Condition $\lambda 2$, or Condition λ for that matter, is no longer necessary. In this example, the choice behavior of the other player is determined by what Manzini and Mariotti (2007) call a rational shortlist method (RSM). In RSM, an individual is sequentially applying different *rationales*, asymmetric binary relations, until only one alternative remains. Formally, a single-valued choice function $C : 2^X \to X$ is an RSM whenever there are two (or more) asymmetric binary relations $P_1, P_2 \subseteq X \times X$, such that[9]

$$C(A) = \max\{\max\{A; P_1\}; P_2\} \text{ for all } A \subseteq X.$$

Notice that *Lexicographic Preferences* (Mas-Colell et al. 1995, p. 46) can be obtained as a special case of RSM when the rationales are chosen in an appropriate way.

[9] The operator *max* is defined as $\max\{A; P\} = \{a \in A \mid cPa \text{ for no } c \in A\}$. An interesting question for future research is whether a similar characterization, as in Moore and Repullo (1990), can be given when all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$ are RSMs (Property α would not hold). This could possibly be based on some kind of *augmented revelation mechanism*, where each player is asked to announce both (or all) asymmetric relations (P_1, P_2).

		Player 2	
		s_2	$\widehat{s}_2$
Player 1	s_1	x	x
	$\widehat{s}_1$	y	z

Fig. 1 A game form that implements f

Example 4 [10] Assume that $X = \{x, y, z\}$ and $I = \{1, 2\}$. Player 1 has only one possible pattern of behavior. The player will act according to an RSM where the primary rationale is $P_1 = \{(z, x)\}$ and the secondary rationale is $P_2 = \{(x, y), (y, z)\}$. The choice behavior that this RSM generates is $C_1(X) = \{y\}$, $C_1(x, y) = \{x\}$, $C_1(x, z) = \{z\}$, $C_1(y, z) = \{y\}$, and hence it does not satisfy Property α. The choice of player 2, on the other hand, may be generated by either a preference relation $P_2 = y \succ z \succ x$ or by a preference relation $P_2' = x \succ z \succ y$. We will denote the corresponding choice functions by C_2 and C_2' respectively. Now, let $\mathscr{C} = \{C_1\} \times \{C_2, C_2'\}$ and define a SCC f by the rule $f(C_1, C_2) = \{x\}$ and $f(C_1, C_2') = \{x, z\}$. Condition λ, which is necessary for Condition $\lambda 2$ to hold, would suggest that f is not implementable in BehNE, because it is impossible to find a set Y that satisfies (iii) of Condition λ. This set should be either $X = \{x, y, z\}$ or $\{x, z\}$, since it must include the range of f, that is $f(\mathscr{C}) = \{x, z\} \subseteq Y$, and neither one satisfies (iii) of Condition λ. To see this, notice that $C_1(X) \cap C_2(X) = \{y\} \notin f(C_1, C_2)$ and $C_1(\{x, z\}) \cap C_2(\{x, z\}) = \{z\} \notin f(C_1, C_2)$. Still, f can be implemented in BehNE by the game form in Fig. 1.

The only BehNE of this game under $\boldsymbol{C} = (C_1, C_2)$ is $\boldsymbol{s} = (s_1, s_2)$, and, there are two BehNE under $\boldsymbol{C} = (C_1, C_2')$, $\boldsymbol{s} = (s_1, s_2)$ and $\boldsymbol{s} = (\widehat{s}_1, \widehat{s}_2)$. □

5 Testing implementability

The usefulness of Condition μ in Moore and Repullo (1990), which is a bit unintuitive, was not truly established until the work of Sjöström (1991).[11] His paper gave an algorithm to test the condition in a constructive way. The algorithm is based on the fact that one can construct a class of test sets that satisfy Condition μ exactly in case SCC is implementable. Unfortunately, for a direct generalization of this algorithm we would need both Property α and Property γ. That is, we would need normality. So even though the characterization of Moore and Repullo (1990) requires only Property

[10] Similar to Example 1, the fact that there are only two players is not critical. We could simply add "dummy" players identical to player 1.

[11] I am indebted to an anonymous referee for pointing out that the algorithm of Sjöström (1991) does not necessarily work if we assume only Property α.

α, efficient verification may sometimes also require Property γ. We will now explain why. For the rest of this section, consider SCC f fixed.

Assume that all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$ satisfy Property α. If there is a profile of choice functions $\boldsymbol{C} \in \mathscr{C}$, such that $\boldsymbol{C}_i(X) = x$ for all $i \in I$ and $x \notin f(\boldsymbol{C})$, then x cannot be used in any game form that implements f. In any game form $g : S \to X$, a strategy profile $\boldsymbol{s} \in S$, such that $g(\boldsymbol{s}) = x$, would be a bad BehNE by Property α. Delete all such alternatives from X to obtain the set $Y^1 \subseteq X$. Since some alternatives were deleted, there may now be new alternatives that cannot be used. This is why we have to continue iteratively

$$Y^1 \supseteq Y^2 \supseteq Y^3, \ldots$$

until $Y^\infty = \bigcap Y^m$ is reached. Let us denote the limit set by Y^*. This set always exists, even though we may not be able to construct it algorithmically when X is infinite. If we should have $Y^* = \emptyset$, then f is clearly not implementable.

Notice that the set Y^* satisfies item (iii) of Condition λ by construction. In the algorithm of Sjöström (1991), Y^* is used as a test set for Y.[12] An important thing to note is that if Property α does not hold, then nothing guarantees that an alternative x, unanimously chosen from X, is a bad BehNE in all game forms where it is used. So, there is no a priori reason to delete it. In fact, there may no longer be a unique maximal set Y^*, in the sense of inclusion, that is unproblematic from a unanimity point of view. To see this, note that when Property α does not hold, we may be able to delete something other than x from X to guarantee that some player no longer chooses it.

Assume then, that all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$ satisfy Property γ. For all $i \in I, \boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$, let $R_i^0(x, \boldsymbol{C})$ be the maximal subset of Y^*, such that $x \in \boldsymbol{C}_i(R_i^0(x, \boldsymbol{C}))$. This is the most natural generalization of a *lower contour set* into a choice function setting. If $R_i^0(x, \boldsymbol{C})$ does not exist for some $i \in I, \boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$, then f obviously cannot be implemented in BehNE. The next Lemma verifies that this concept is well defined.

Lemma 1 *Let C be a choice function that satisfies Property γ and assume that $x \in A \subseteq X$. The maximal subset $B \subseteq A$, such that $x \in C(B)$, is unique whenever it exists.*

Proof Notice first that since deadlocks are allowed, x need not be chosen from any set. For the sake of contradiction, assume there are two maximal subsets $B, B' \subseteq A$, such that $x \in C(B)$ and $x \in C(B')$. Since both B and B' are maximal, we cannot have $B \subseteq B'$ or $B' \subseteq B$. But Property γ implies that $x \in C(B \cup B')$ must be the case, so neither B nor B' can in fact be maximal. This is a contradiction. Therefore, the maximal subset from which an alternative is chosen must be unique whenever it exists. □

If $R_i^0(x, \boldsymbol{C})$ together with Y^* do not satisfy item (ii) of Condition λ, then there must be an alternative $y \in R_i^0(x, \boldsymbol{C}) \subseteq Y^*$ and a profile of choice functions $\boldsymbol{C}^* \in \mathscr{C}$, such that

[12] See Condition λ.

$$y \in \boldsymbol{C}_i^*(R_i{}^0(x, \boldsymbol{C})) \bigcap \left[\bigcap_{j \neq i} \boldsymbol{C}_j^*(Y) \right] \text{ and } y \notin f(\boldsymbol{C}^*).$$

Delete all such alternatives from $R_i{}^0(x, \boldsymbol{C})$ to obtain the set $R_i{}^1(x, \boldsymbol{C})$. Since some alternatives were deleted, there may now be new alternatives that do not satisfy item (ii) of Condition λ. Again, we have to continue iteratively

$$R_i{}^0(x, \boldsymbol{C}) \supseteq R_i{}^1(x, \boldsymbol{C}) \supseteq R_i{}^2(x, \boldsymbol{C}), \ldots$$

until $R_i{}^\infty(x, \boldsymbol{C}) = \bigcap R_i^m(x, \boldsymbol{C})$ is reached. Let us denote the limit set by $R_i^*(x, \boldsymbol{C})$. In the algorithm of Sjöström (1991), this set is used as a test set for $R_i(x, \mathbf{C})$. As in the case of Y^*, if Property γ does not hold, we may have several different candidates for $R_i{}^0(x, \boldsymbol{C})$—as should be obvious by the Lemma above. The observations in this section gives us the following theorem.

Theorem 5 *Assume that every choice function in $\bigcup_{i \in I} \mathscr{C}_i$ satisfies Property α and Property γ. That is, all choice functions in the domain are normal. Assume that $n \geq 3$ and $x \in R_i^*(x, \boldsymbol{C})$ for all $i \in I, \boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$. An SCC f is implementable in BehNE if and only if the choice $Y = Y^*$ and $R_i(x, \boldsymbol{C}) = R_i^*(x, \boldsymbol{C})$ for all $i \in I, \boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$ satisfies Condition λ.*

Proof [13] Notice first that the sets Y^* and $R_i^*(x, \boldsymbol{C})$ will satisfy items (ii) and (iii) of Condition λ by construction. Furthermore, and again simply by construction, these are the maximal sets that can satisfy them. But then, if these sets do not satisfy item (i) of Condition λ, no subset of them can do so by Property α either. This verifies our claim. □

It is important to understand that the verification of Condition λ is not difficult if X is finite. We can simply use direct search, that is, we can try all possible combinations of sets. The main point is that if all choice functions in the domain are normal, then we can construct *one* class of test sets that satisfy Condition λ exactly in case SCC f is implementable in BehNE. This is what makes the algorithm of Sjöström (1991) so efficient. In contrast, when normality does not hold, whether because Property α does not hold or because Property γ does not hold, there may no longer be only one class of test sets. As we have seen, several sets can be chosen as Y^* or $R_i^*(x, \mathbf{C})$. Since Condition λ ties these sets together, we have to test all possible combinations. This will rapidly expand the number of classes we have to run through and make the algorithm much more inefficient.

6 Concluding remarks

We have shown that the results of implementation theory do not require the assumption of rationality per se. Rather, what is important, is that the assumption of rationality

[13] Notice that Property γ is used only implicitly in this proof to guarantee that there is a unique set we can choose as $R_i^*(x, \boldsymbol{C})$.

render certain types of framing impossible (Property α does not hold). However, if we do not assume full rationality, then the verification of the characterizing condition may sometimes be more challenging.

It should not come as a surprise that some form of regularity in behavior is needed if results originally designed to hold when all players are fully rational are generalized. Still, it is an important topic for future research to find out what can be implemented when this regularity—Property α—does not hold. The behavior that we observe in reality does not always satisfy this property, an interesting example being the RSM of Manzini and Mariotti (2007). In RSM, an individual is sequentially applying different rationales until only one alternative remains—a heuristic that we all use, consciously or unconsciously, in multi-criteria decision making. But we would need a new canonical mechanism to give a characterization that does not use Property α. The characterization of Moore and Repullo (1990) is based on the fact that one can construct an implementing game form by using the class of sets in Condition λ. We learned in Example 4 that this cannot necessarily be done if Property α does not hold.

As a final point we want to stress that there is nothing wrong in the result of Hurwicz (1986). If players behave according to GNE, then so be it. This is as good behavioral assumption as any. The main idea was to demonstrate that there is a more general solution concept, that of BehNE, which can be applied to a wider class of implementation problems. This follows directly from the fact that BehNE coincides with GNE if all choice functions in the domain of a SCC are normal.

Appendix

Proof of Theorem 1 (*Sufficiency*) Let the strategy set of player i be

$$S_i = \{(\boldsymbol{C}, x, y, n) \in \mathscr{C} \times X \times Y \times \mathbb{N} \mid x \in f(\boldsymbol{C})\},$$

and, define the mechanism $g\colon S \to X$ by the following three mutually exhaustive rules:

(1) If there is a strategy $(\boldsymbol{C}, x, y, n) \in \mathscr{C} \times X \times Y \times \mathbb{N}$, such that $s_i = (\boldsymbol{C}, x, y, n)$ for all $i \in I$, then $g(\boldsymbol{s}) = x$.

(2) If there is a strategy $(\boldsymbol{C}, x, y, n) \in \mathscr{C} \times X \times Y \times \mathbb{N}$ and a player $i \in I$, such that $s_j = (\boldsymbol{C}, x, y, n)$ for all $j \neq i$ and $s_i = (\boldsymbol{C}', x', y', n') \neq (\boldsymbol{C}, x, y, n)$, then

$$g(\boldsymbol{s}) = \begin{cases} y' & \text{if } y' \in R_i(x, \boldsymbol{C}), \\ x & \text{otherwise.} \end{cases}$$

(3) If neither (1) nor (2) applies, and the strategy of player i is denoted by $s_i = (\boldsymbol{C}^i, x^i, y^i, n^i)$, then $g(\boldsymbol{s}) = y^k$ for $k = \min\{j \in I \mid n_j \geq n_i \text{ for all } i \in I\}$.[14]

We shall prove the claim by showing that $f(\boldsymbol{C}) = \{g(\boldsymbol{s}) \mid \boldsymbol{s} \in BehNE\,[\Gamma(\boldsymbol{C})]\}$ for all $\boldsymbol{C} \in \mathscr{C}$. First, we verify that $f(\boldsymbol{C}) \subseteq \{g(\boldsymbol{s}) \mid \boldsymbol{s} \in BehNE\,[\Gamma(\boldsymbol{C})]\}$ for all $\boldsymbol{C} \in \mathscr{C}$.

[14] The index selection method is simply a tie breaking rule.

Let $x \in f(\boldsymbol{C})$ and consider the strategy profile $\boldsymbol{s} \in S$, such that $s_i = (\boldsymbol{C}, x, x, 0)$ for all $i \in I$. By the definition of g player i can obtain exactly the set $R_i(x, \boldsymbol{C})$ through a unilateral deviation from the strategy profile $\boldsymbol{s}$. This strategy profile is then a BehNE of the game $\Gamma(\boldsymbol{C})$ and $g(\boldsymbol{s}) = x$. Since this holds for all $\boldsymbol{C} \in \mathscr{C}$ and $x \in f(\boldsymbol{C})$, we have established the claim.

Second, we will verify that $\{g(\boldsymbol{s}) \mid \boldsymbol{s} \in BehNE[\Gamma(\boldsymbol{C}^*)]\} \subseteq f(\boldsymbol{C}^*)$ for all $\boldsymbol{C}^* \in \mathscr{C}$.[15] Pick $\boldsymbol{s} \in BehNE[\Gamma(\boldsymbol{C}^*)]$. There are three cases to consider, as the BehNE can come from any one of cases (1), (2) or (3). If (1) applies, and $\boldsymbol{C}^*$ is the true profile of choice functions, then $g(\boldsymbol{s}) = x \in \bigcap_{i \in I} \boldsymbol{C}_i^*(R_i(x, \boldsymbol{C}))$, so that $g(\boldsymbol{s}) \in f(\boldsymbol{C}^*)$ by item (i) of Condition λ as required. If (2) applies, and $\boldsymbol{C}^*$ is the true profile of choice functions, then $g(\boldsymbol{s}) \in \boldsymbol{C}_i^*(R_i(x, \boldsymbol{C})) \bigcap \left[\bigcap_{j \neq i} \boldsymbol{C}_j^*(Y)\right]$ for some player i,[16] so that $g(\boldsymbol{s}) \in f(\boldsymbol{C}^*)$ by item (ii) of Condition λ as required. Finally, if (3) applies, and $\boldsymbol{C}^*$ is the true profile of choice functions, then $g(\boldsymbol{s}) \in \bigcap_{i \in I} \boldsymbol{C}_i^*(Y)$,[17] so that $g(\boldsymbol{s}) \in f(\boldsymbol{C}^*)$ by item (iii) of Condition λ. Since this holds for all $\boldsymbol{s} \in BehNE[\Gamma(\boldsymbol{C}^*)]$, no matter which one of cases (1), (2) or (3) gives rise to the BehNE, we have established the claim. This completes the proof of sufficiency. □

Proof of Theorem 2 (*Necessity*) Assume that SCC f is implemented in BehNE by the game form $G = (S, g)$. First, define the set Y by the rule

$$Y \equiv \{x \in X \mid x = g(\boldsymbol{s}) \text{ for some } \boldsymbol{s} \in S\}.$$

Second, choose a profile of choice functions $\boldsymbol{C} \in \mathscr{C}$ and an outcome $x \in f(\boldsymbol{C})$. Take any behavioral Nash equilibrium $\boldsymbol{s}(x, \boldsymbol{C})$ of the game $\Gamma(\boldsymbol{C})$, such that $g(\boldsymbol{s}(x, \boldsymbol{C})) = x$. Let $R_i(x, \boldsymbol{C})$ be the set of all alternatives that player i can obtain by unilaterally deviating from $\boldsymbol{s}(x, \boldsymbol{C})$, that is

$$R_i(x, \boldsymbol{C}) \equiv \{y \in Y \mid y = g(s_i, \boldsymbol{s}(x, \boldsymbol{C})_{-i}) \text{ for some } s_i \in S_i\}.$$

We will prove that these sets satisfy items (i),(ii) and (iii) of Condition λ as a direct consequence of the fact that G implements f in BehNE. Assume that the left side of implication in (i) is satisfied, that is, all players i choose x from the set $R_i(x, \boldsymbol{C})$ when the true profile of choice functions is $\boldsymbol{C}^*$. By definition, $\boldsymbol{s}(x, \boldsymbol{C})$ is then also BehNE in the game $\Gamma(\boldsymbol{C}^*)$. Since f is implemented by the game form G, this implies $x \in f(\boldsymbol{C}^*)$ as required in (i). Assume then, that the left side of the implication in (ii) is satisfied. By definition, there is a strategy $s_i \in S_i$, such that $y = g(s_i, \boldsymbol{s}(x, \boldsymbol{C})_{-i})$. Since alternative y is the choice from Y by all players $j \in I \setminus \{i\}$, strategy profile $(s_i, \boldsymbol{s}(x, \boldsymbol{C})_{-i})$ must be a BehNE of the game $\Gamma(\boldsymbol{C}^*)$ by Property α. This implies $y \in f(\boldsymbol{C}^*)$, as required in (ii). Finally, assume that the left side of the implication in (iii) holds. Now, there must be a strategy profile $\boldsymbol{s} \in S$, such that $z = g(\boldsymbol{s})$. Again, by Property α, this strategy

[15] We are here using $\boldsymbol{C}^*$ instead of $\boldsymbol{C}$ because we have to make a distinction between the *true* profile of choice functions and the choice functions that are given as part of the strategies.

[16] Every player $j \neq i$ could unilaterally deviate to any alternative in Y by announcing a high enough integer.

[17] In this case, every player can deviate to any alternative in Y.

profile must be a BehNE of the game $\Gamma(\boldsymbol{C}^{*})$, and so we have $z \in f(\boldsymbol{C}^{*})$ as required in (iii).[18] This completes the proof of necessity. $\square$

Acknowledgments I thank Hannu Vartiainen and Hannu Salonen for useful discussions and generous help. I also thank the participants of the weekly seminar of Public Choice Research Centre, especially Hannu Nurmi. The comments of two anonymous referees and an associate editor have been more than helpful and greatly improved the quality of this article.

References

Bar-Hillel, M., & Margalit, A. (1988). How vicious are cycles of intransitive choice? *Theory and Decision, 24*, 119–145.

Chernoff, J. S. (1954). Rational selection of decision functions. *Econometrica, 22*, 423–443.

Dutta, B., & Sen, A. (1991). A necessary and sufficient condition for two-person Nash implementation. *Review of Economic Studies, 58*, 121–128.

Hurwicz, L. (1986). On the implementation of social choice rules in irrational societies. In W. Heller, R. Starrett, & D. Starr (Eds.), *Social choice and public decision making. Essays in Honor of Kenneth J. Arrow* (Vol. I). Cambridge University Press, Cambridge.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341–350.

Manzini, P., & Mariotti, M. (2007). Sequentially rationalizable choice. *The American Economic Review, 97*, 1824–1839.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. USA: Oxford University Press.

Maskin, E. (1977). *Nash equilibrium and welfare optimality*. Cambridge, Mimeo: Department of Economics, MIT.

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies, 66*, 23–38.

Matsushima, H. (2008a). Behavioral aspects of implementation theory. *Economics Letters, 100*, 161–164.

Matsushima, H. (2008b). Role of honesty in full implementation. *Journal of Economic Theory, 139*, 353–359.

Moore, J., & Repullo, R. (1990). Nash implementation—a full characterization. *Econometrica, 58*, 1083–1099.

Moulin, H. (1985). Choice functions over a finite set: A summary. *Social Choice and Welfare, 2*, 147–160.

Nash, J. F. (1950). The bargaining problem. *Econometrica, 18*, 155–162.

Plott, C. R. (1976). Social choice theory: An overview and interpretation. *American Journal of Political Science, 20*, 511–596.

Ray, K. T. (2010) Nash implementation under irrational preferences. Working Paper. Washington DC, Mimeo: McDonough School of Business.

Rubinstein, A., & Salant, Y. (2008). (A, f): Choice with frames. *The Review of Economic Studies, 75*, 1287–1296.

Samuelson, P. (1938). A note on the pure theory of consumer's behavior. *Economica, 5*, 61–71.

Sen, A. (1977). Social choice theory: A re-examination. *Econometrica, 45*, 53–89.

Sen, A. (1970). Choice functions and revealed preference. *The Review of Economic Studies, 38*, 307–317.

Sjöström, T. (1991). On the necessary and sufficient conditions for Nash implementation. *Social Choice and Welfare, 8*, 333–340.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*, 31–48.

[18] Notice how we are using Property α here: If alternative x is chosen from the set Y, then it must be chosen from any row in the game form that contains it.

ARTICLE 4

On the Behavioral Boundaries of the Revelation Principle

On the Behavioral Boundaries of the Revelation Principle

VILLE KORPELA [*] [†]
Public Choice Research Centre (PCRC),
University of Turku

18.12.2009
Last Revision 07.06.2011

Abstract

The revelation principle is the foundation stone of mechanism design. This principle states that the search for incentive compatible mechanisms can be restricted in the set of direct mechanisms without loss of generality. In this paper we ask: What are the minimal assumptions on behavior for the revelation principle to still work? It turns out that a well-known property, called Property α, is central.
JEL Classification: D03, D71, D78
Keywords: Property α; Revelation Principle; Truthful Implementation

1. INTRODUCTION

The revelation principle forms the foundation of mechanism design by considerably simplifying the search for incentive compatible mechanisms (see Dasgupta et al., 1979; Harris and Townsend, 1981; and Myerson, 1979, for example). This principle states that, assuming truthful implementation is considered satisfactory, one can focus on direct mechanisms only. That is,

[*] E-mail: ville.korpela@utu.fi.
[†] Many thanks to Hannu Vartiainen for useful discussions.

one can restrict attention in the set of mechanisms that simply ask individuals to truthfully reveal their type.

Recently, there has been a growing trend in implementation literature to incorporate more behavioral assumption into the theory (see Matsushima, 2008a; and Matsushima, 2008b, for example). Since the revelation principle was originally formulated in a fully rational environment,[1] it is important to find out the exact behavioral boundaries for this central principle of mechanism design. If it turns out that revelation principle does not hold for the most natural types of individual behavior, then a lot more attention has to be paid towards indirect mechanisms in the future.

This topic has been previously studied by Saran (2011) in the case of Bayesian-Nash equilibrium. In contrast to Saran, our purpose is to solve the question for dominant strategy equilibrium. Even though there are some parallels between implementation in Bayesian-Nash equilibrium and implementation in dominant strategy equilibrium, the latter being a robust version of the former (see Bergemann and Morris, 2005), the case of dominant strategy equilibrium still demands a treatment of its own for at least two reasons. First, dominat strategy equilibrium is the most natural solution concept used in a mechanism design problem, and second, the results are much more transparent, an hence easier to understand. Moreover, it is not immediate that the connection between these two implementation forms remain when rationality is not assumed, even though the central position of Property α in both papers suggest that it does.

The rest of this short paper is organized as follows. In Sect. 2 we fix notation and define basic concepts. We also present the framework of Hurwicz (1986) for incorporating more behavioral assumption into implementation theory. Sect. 3 contains all the main results. We show that a well-known property, called Property α, is a necessary and sufficient condition for the revelation principle to work when no reference is made to the social choice function (a

[1] By rationality we mean that choice behavior is generated by a complete and transitive preference relation.

condition imposed only upon the domain of behavior). We give also a general characterization, which is a simple generalization of strategy-proofness that we dub behavioral strategy-proofness. Sect. 4 concludes.

2. NOTATION AND PRELIMINARIES

Let $I = \{1, \ldots, n\}$ be the set of agents and A the set of (social) alternatives.[2] A typical element of I is denoted by i and a typical element of A is denoted by a, b or c, and so forth. A *choice function* is a mapping $C : 2^A \setminus \emptyset \to 2^A$, such that $C(B) \subseteq B$ for all $B \in 2^A \setminus \emptyset$.[3] The case $C(B) = \emptyset$ is not precluded. We interpret this by saying that an agent refuses to make any choice at all, that is, there is a *deadlock*. A choice function C is generated by the complete and transitive binary relation $R \subseteq A \times A$ if and only if

$$C(B) = \{a \in B \mid a\, R\, b \text{ for all } b \in B\} \text{ for all } B \in 2^A.$$

If C is generated by R, then we simply write $C = R$.[4] A few well-known properties of choice functions are needed.

Property α : For all $B, D \subseteq A$, if $a \in B \subseteq D$ and $a \in C(D)$, then it must also be true that $a \in C(B)$.

Property β : For all $B, D \subseteq A$, if $a, b \in C(B)$, $B \subseteq D$ and $b \in C(D)$, then it must also be true that $a \in C(D)$.

Basically, Property α states that a chosen alternative a must still be choosable when the choice set is reduced by dropping other alternatives. Property β, on the other hand, states that if two alternatives a and b are "*inseparable*" in a smaller set, then these alternatives cannot be "*separated*" in any larger choice set either. Both of these axioms are consistency requirements, and taken together, they imply that choice behavior is rationalizable.[5] In a sense,

---

[2] We do not rule out the case $n = 1$.

[3] 2^A is the *power set* of A, that is, $2^A = \{B \mid B \subseteq A\}$.

[4] With a little abuse of notation

[5] Choice function C is *rationalizable* when a complete and transitive binary relation exists that generates it. See Kreps (1988) and Sen (1977), and the references given there, for a full treatment.

then, rationality can be interpreted as a very strict form of consistency in behavior.

The set of all possible choice functions over A is denoted by $\mathscr{C}_A$. A *Social Choice Correspondence* (SCC) is any correspondence $f : \mathscr{C} \twoheadrightarrow A$, such that $\mathscr{C} \subseteq \bigtimes_{i\in I} \mathscr{C}_i$ and $\mathscr{C}_i \subseteq \mathscr{C}_A$ for all $i \in I$. If f is in fact a function, then we call it a *Social Choice Function* (SCF) and denote $f : \mathscr{C} \to A$. We interpret this by saying that social choice is based on behavioral patterns. A typical element of $\mathscr{C}$ is denoted by $\boldsymbol{C}$ and the choice function of agent i, in this profile, by $\boldsymbol{C}_i$. As usual, $\boldsymbol{C}_{-i}$ denotes a profile of choice functions that specifies a choice function for all agents except i. If all choice functions that are admissible are generated by a complete and transitive binary relation (the standard implementation setting), then we simply write $\mathscr{C}_A = \mathscr{R}_A$, $\mathscr{C}_i = \mathscr{R}_i$, $\mathscr{C} = \mathscr{R}$, $\boldsymbol{C} = \boldsymbol{R}$ and $\boldsymbol{C}_{-i} = \boldsymbol{R}_{-i}$ respectively. Finally, for any function $q : X \to Y$ and any set $Z \subseteq X$, we use the standard convention $q(Z) = \{q(z) \mid z \in Z\} \subseteq Y$.

A tuple $G = (S, g)$, where $S = S_1 \times \cdots \times S_n$ is the *set of strategy profiles* and $g : S \to A$ is the *outcome function*, is called a *mechanism*. We sometimes identify the mechanism by giving only the outcome function g. For a fixed profile of preference relations $\boldsymbol{R} = (R_1, \ldots, R_n)$, this mechanism defines a game $\mathit{\Gamma}(\boldsymbol{R}) = (G; \boldsymbol{R})$. Strategy profile $\boldsymbol{s}$ is a *dominant strategy equilibrium* of this game if and only if $g(\boldsymbol{s}_i, \widehat{\boldsymbol{s}}_{-i}) \boldsymbol{R}_i\ g(\widehat{\boldsymbol{s}})$ for all $\widehat{\boldsymbol{s}} \in S$ and all $i \in I$. The set of all dominant strategy equilibria of the game $\mathit{\Gamma}(\boldsymbol{R})$ is denoted by $\mathrm{DSE}\big[\mathit{\Gamma}(\boldsymbol{R})\big] \subseteq S$. We say that mechanism g *implements an SCC f in dominant strategies* if

$$g\big(\mathrm{DSE}\big[\mathit{\Gamma}(\boldsymbol{R}\,)\big]\big) = f(\boldsymbol{R}) \;\text{ for all }\; \boldsymbol{R} \in \mathscr{R}. \tag{1}$$

That is, if the outcomes of dominant strategy equilibria coincide exactly with the SCC f. If a mechanism that implements f in dominant strategies exists, then we say that f is *dominant strategy implementable.* A mechanism h, that has the set of preferences as the set of strategies ($S_i = \mathscr{R}_i$), is called a *direct mechanism.* If

$$\Big[\boldsymbol{R} \in \mathrm{DSE}\big[G(h; \boldsymbol{R})\big] \text{ and } h(\boldsymbol{R}) \in f(\boldsymbol{R})\Big] \text{ for all } \boldsymbol{R} \in \mathscr{R}, \tag{2}$$

then we say that h *implements* f *truthfully*.[6] Note that all dominant strategy equilibria of h do not have to coincide with f, only the truth-telling one.

3. MAIN RESULT

To get started, we state and prove the influential Revelation Principle. Our proof follows closely that given in Dasgupta et al. (1979).[7] Notice that if the designer of a social choice mechanism is willing to implement the SCC f, then to be consistent with the idea that all alternatives in the range of f are socially acceptable, the designer should also be willing to implement any selection from f. That is, the designer should be willing to implement any function $s : \mathscr{R} \to A$, such that $s(\boldsymbol{R}) \in f(\boldsymbol{R})$ for all $\boldsymbol{R} \in \mathscr{R}$. The fact that SCC f is implementable does not, of course, imply that all selections from it are implementable. We are only saying that if some selection would be implementable, then it should not make any difference whether the designer implements the SCC f or the selection from it.

Theorem 1 (*The Revelation Principle*). If SCC $f : \mathscr{R} \twoheadrightarrow A$ is implementable in dominant strategies, then at least one selection from it is implementable truthfully.

REMARK I. *Truthful implementation is satisfactory when truth-telling is a focal strategy. If one cannot gain anything by lying, it might be reasonable to assume that one does not.*[8]

Proof. Let $G = (S, g)$ be a mechanism that implements f in dominant strategies and let $s^* : \mathscr{R} \to S$ be a selection that associates every preference profile $\boldsymbol{R} \in \mathscr{R}$ with some dominant strategy equilibrium of the mechansim G. That is, $s^*(\boldsymbol{R}) \in \mathrm{DSE}\big[\mathit{\Gamma}(\boldsymbol{R}\,)\big]$ for all $\boldsymbol{R} \in \mathscr{R}$. Since dominant strategies of agent i do not depend on the strategies chosen by others, there must exist at least one selection that satisfies $s^*(\boldsymbol{R}) = (s_i^*(\boldsymbol{R}_i), s_{-i}^*(\boldsymbol{R}_{-i}))$ for all $\boldsymbol{R} \in \mathscr{R}$ and all $i \in I$. Use this selection to define a direct mechanism $h : \mathscr{R} \to A$ by

[6] In this case, h is sometimes called *straightforward*.

[7] See also Repullo (1985).

[8] Recent experimental results indicate that truth-telling is not always focal (see Saijo, 2005, for example).

the rule:

$$h(\boldsymbol{R}) = g(s^*(\boldsymbol{R})) \text{ for all } \boldsymbol{R} \in \mathscr{R}.$$

It is sufficient to show that h implements f truthfully. To this end, fix $\boldsymbol{R} \in \mathscr{R}$ and $i \in I$. By the definition of dominant strategy equilibrium we have

$$g(s_i^*(\boldsymbol{R}_i), s_{-i}^*(\boldsymbol{R}'_{-i}))\boldsymbol{R}_i g(s_i^*(\boldsymbol{R}'_i), s_{-i}^*(\boldsymbol{R}'_{-i})) \text{ for all } \boldsymbol{R}' \in \mathscr{R},$$

which, by the definition of h, can be re-written as

$$h(\boldsymbol{R}_i, \boldsymbol{R}'_{-i})\boldsymbol{R}_i h(\boldsymbol{R}'_i, \boldsymbol{R}'_{-i}) \text{ for all } \boldsymbol{R}' \in \mathscr{R}.$$

This completes the proof, since truth-telling is a dominant strategy equilibrium of h. ■

To formulate the Revelation Principle in a more general behavioral setting, we must first generalize the idea of dominant strategy equilibrium. Notice that a mechanism $G = (S, g)$ will define a game $\Gamma(\boldsymbol{C})$ also for a fixed profile of choice functions $\boldsymbol{C} = (C_1, \ldots, C_n)$. The specification of choice functions will tell us how agents act when they face any row of the mechanism.

Definition 1. Strategy profile $\boldsymbol{s}$ is a *Strong Behavioral Equilibrium* (SBE) of the game $\Gamma(\boldsymbol{C})$, induced by the mechanism $G = (S, g)$, if and only if

$$g(\boldsymbol{s}_i, \boldsymbol{s}'_{-i}) \in \boldsymbol{C}_i(g(S_i, \boldsymbol{s}'_{-i})) \text{ for all } \boldsymbol{s}'_{-i} \in S_{-i} \text{ and all } i \in I.$$

This definition is a straightforward generalization of the idea behind dominant strategy equilibrium: Agent i is willing to play the strategy $\boldsymbol{s}_i$ despite the strategies chosen by others. Both implementation formulas of section 2, formula (1) and (2), can be directly generalized for this solution concept (set DSE=SBE, $\mathscr{R} = \mathscr{C}$ and $\boldsymbol{R} = \boldsymbol{C}$). The main result of this paper can now be stated, and follows below.

Theorem 2. Assume that all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$ satisfy Property α. If SCC $f : \mathscr{C} \twoheadrightarrow A$ is implementable in SBE, then at least one selection from it is implementable truthfully.

REMARK II. *If all choice functions in $\bigcup_{i \in I} \mathscr{C}_i$ are rationalizable, then we are back to Theorem 1. Here, we do not assume Property β, only Property α.*[9]

Proof. Let $G = (S, g)$ be a mechanism that implements f in SBE and let $s^* : \mathscr{C} \to A$ be a selection that associates every profile of choice functions $\boldsymbol{C} \in \mathscr{C}$ with some SBE of the mechanism G. That is, $s^*(\boldsymbol{C}) \in \mathrm{SBE}\big[\Gamma(\boldsymbol{C})\big]$ for all $\boldsymbol{C} \in \mathscr{C}$. For the same reason as in the case of dominant strategies, there must be at least one selection that satisfies $s^*(\boldsymbol{C}_i, \boldsymbol{C}_{-i}) = (s_i^*(\boldsymbol{C}_i), s_{-i}^*(\boldsymbol{C}_{-i}))$ for all $\boldsymbol{C} \in \mathscr{C}$ and all $i \in I$. Use this selection to define a direct mechanism $h : \mathscr{C} \to A$ by the rule:

$$h(\boldsymbol{C}) = g(s^*(\boldsymbol{C})) \text{ for all } \boldsymbol{C} \in \mathscr{C}.$$

Again, it is sufficient to show that h implements f truthfully. So fix $\boldsymbol{C} \in \mathscr{C}$ and $i \in I$. By the definition of SBE we have

$$g(s_i^*(\boldsymbol{C}_i), \boldsymbol{s}_{-i}) \in \boldsymbol{C}_i(g(S_i, \boldsymbol{s}_{-i})) \text{ for all } \boldsymbol{s}_{-i} \in S_{-i} \text{ and all } i \in I.$$

Since $g(s_i^*(\mathscr{C}_i), \boldsymbol{s}_{-i}) \subseteq g(S_i, \boldsymbol{s}_{-i})$ for all $\boldsymbol{s}_{-i} \in S_{-i}$, and since $\boldsymbol{C}_i$ satisfies Property α by assumption, this implies that for all $i \in I$ we must have

$$g(s_i^*(\boldsymbol{C}_i), s_{-i}^*(\boldsymbol{C}'_{-i})) \in \boldsymbol{C}_i(g(s_i^*(\mathscr{C}_i), s_{-i}^*(\boldsymbol{C}'_{-i}))) \text{ for all } \boldsymbol{C}'_{-i} \in \mathscr{C}_{-i},$$

which, by the definition of h, can be re-written as

$$h(\boldsymbol{C}_i, \boldsymbol{C}'_{-i}) \in \boldsymbol{C}_i(h(\mathscr{C}_i, \boldsymbol{C}'_{-i})) \text{ for all } \boldsymbol{C}'_{-i} \in \mathscr{C}_{-i}.$$

This completes the proof, since truth-telling is an SBE of h. ■

In the next example we show that when Property α does not hold, Revelation Principle may not hold either. The second example is an application of Theorem 2.

[9] Deadlocks are now allowed as long as Property α holds. Notice how this comes about in the proof. If mechanism $G = (S, g)$ implements f in SBE, then deadlocks cannot occur when any agent is choosing from any row by Definition 1, so deadlocks cannot occur in mechanism h either by Property α.

Example 1 (Seller Choosing a Product Variety)**.** Let us assume that a seller has three products to offer for a potential customer, that is $A = \{a, b, c\}$. Moreover, let us denote the profit from a sale of a, b and c by $\pi(a)$, $\pi(b)$ and $\pi(c)$ respectively, and the price to customers by $p(a), p(b)$ and $p(c)$ respectively. Assume that, for whatever reason, we have

$$\pi(b) > \pi(a) > \pi(c), \text{ but } p(a) > p(b) > p(c).$$

Furthermore, assume that there are two types of customer, so that $\mathscr{C} = \mathscr{C}_1 = \{C_1, C_2\}$. These customers act according to the following *rule of thumb*: If the seller has only one product to offer, then both types refuse to buy and go elsewhere. Formally, if $B \in 2^A$ is a singleton, then $C_1(B) = C_2(B) = \emptyset$. If more than one product is offered, then type one customer will buy the second-most cheap and type two customer will buy the second-most expensive. Notice that neither C_1 nor C_2 satisfies Property α. The seller would obviously want to implement a sales function $S : \mathscr{C} \to A$, such that $S(C_1) = S(C_2) = b$. This would maximize her profit. This sales function is implementable in SBE. Simply, offer a product variety $\{a, b, c\}$ and let both types of customer self-select. Every customer will then select product b and the seller will maximize the profit by keeping a product variety containing items that are never actually bought. Still, no direct mechanism can implement S truthfully. The range of a direct mechanism would be $\{b\}$, suggesting that this should be the product variety of the seller. But then, all customers would go elsewhere.[10] $\square$

Example 2 (Voting Rule)**.** Let $I = \{1, 2, 3\}$, $A = \{a, b, c\}$ and $\mathscr{C} = \mathscr{P}_A \times \mathscr{P}_A \times \{C\}$. Here $\mathscr{P}_A$ is the set of all linear orderings over A and we define choice function C by the rule: $C(\{a, b\}) = a$, $C(\{b, c\}) = b$, $C(\{a, c\}) = c$ and $C(\{a, b, c\}) = \emptyset$. Choice function C satisfies Property α and we interpret it by saying that agent 3 cannot decide which alternative, a, b or c, is the best. Let $m(P)$ be the best alternative with respect to $P \in \mathscr{P}_A$. Define a

[10] The fact that we allow deadlocks is not critical for this example. Notice that deadlocks can be interpreted here as a participation constraint.

voting rule $f : \mathscr{C} \to A$ by the following condition:

$$f(P_1, P_2, C) = \begin{cases} m(P_1), & \text{if } m(P_1) = m(P_2), \\ C\left(\{m(P_1), m(P_2)\}\right), & \text{otherwise.} \end{cases}$$

That is, an alternative is selected if agent 1 and agent 2 are unanimous, otherwise agent 3 decides. This voting rule cannot be truthfully implemented in SBE. Hence, it cannot be implemented in SBE by any indirect mechanism either, according to Theorem 2. To see this, let $P_1 = a \succ b \succ c$ and $P_2 = c \succ b \succ a$. Since agent 3 will choose b or c from any set of two alternatives that contain c, we have $f(\mathscr{P}_A, P_2, C) = \{b, c\}$. On the other hand $f(P_1, P_2, C) = c$, so that truth-telling is not an SBE of $G = (f, \mathscr{C})$ under the choice function profile $\boldsymbol{C} = (P_1, P_2, C)$. $\square$

It must be emphasized that Example 1 does not in any way prove that Property α is necessary for the Revelation Principle to work. In fact, the Revelation Principle generally holds due to a specific connection between the domain of all admissible profiles of choice functions $\mathscr{C}$ and the SCC $f : \mathscr{C} \twoheadrightarrow A$. The following definition gives us the exact boundaries of the Revelation Principle.

Definition 2. An SCF $f : \mathscr{C} \to A$ is *behaviorally strategy-proof* if and only if

$$f(\boldsymbol{C}) \in \boldsymbol{C}_i(f(\mathscr{C}_i, \boldsymbol{C}_{-i})) \text{ for all } i \in I \text{ and all } \boldsymbol{C} \in \mathscr{C}.$$

It should be obvious that the Revelation Principle will hold for a given SCC if and only if at least one selection from it is behaviorally strategy-proof. Note that behavioral strategy-proofness is *not* a necessary condition for an SCC to be implementable in SBE.[11]. On the other hand, Revelation Principle applies in the standard rational environment simply because strategy-proofness *is* a necessary condition for dominant strategy implementation.[12] Therefore, the implementation approach based on choice functions has a gap

[11]Example 1 verifies this when deadlocks are allowed. A similar example without deadlocks could easily be given.

[12]See Dasgupta et al. (1979) or Repullo (1985). For a lengthier introduction to strategy-proofness, see Barberà (2001).

between full implementation and truthful implementation. Something that we do not see in the standard rational environment.

However, our goal was to find a condition that allow us to use only direct mechanisms, without any reference to the SCC. Example 1 established that this condition is Property α by showing that the Revelation Principle may not hold when Property α does not hold. Of course, whether it does or not, depends on the SCC.

4. CONCLUDING COMMENTS

It has become evident that menu-dependence is a part of human behavior (see Manzini and Mariotti, 2007; and Rubinstein and Salant, 2008, for example). The fact that an item is selected from the menu $M = \{a_1, \ldots, a_m\}$, does not necessarily imply that it will be selected from every sub-menu $M' \subset M$ that contains it. And yet, this is exactly the kind of behavior that we have to rule out (by assuming Property α) to guarantee that the revelation principle still works. Otherwise, mechanism design problems are bound to get much harder.

It appears, then, that a lot more attention has to be paid to indirect mechanisms in the future. This is an almost entirely unexplored territory. Indeed, we do not even have a clear picture about the SCCs that are important in environments that exhibit menu-dependence. It might be that behavioral strategy-proofness is a predominant property of these SCCs. However, all of these question are well beyond the scope of this paper, and therefore left for future research.

References

Barberà, S. (2001), An Introduction to Strategy-Proof Social Choice Functions. *Social Choice and Welfare* 18, 619-653.

Bergemann, D. and Morris, S. (2005), Robust Mechanism Design. *Econometrica* 73, 1771-1813.

Dasgupta, P., Hammond, P. and Maskin, E. (1979), The Implementation of Social Choice Rules – Some General Results on Incentive Compatibility. *The Review of Economic Studies* 46, 185-216.

Harris, M. and Townsend, R. (1981), Resource Allocation with Asymmetric Information. *Econometrica* 49, 33-64.

Hurwicz, L. (1986), On the Implementation of Social Choice Rules in Irrational Societies. In *Social Choice and Public Decision Making, Essays in Honor of Kenneth J. Arrow Volume I*, ed. by W. Heller, R. Starr and D. Starrett. Cambridge University Press, USA.

Kreps, D. (1988), *Notes on the Theory of Choice.* Undergroung Classics in Economics, Westview Press, USA.

Manzini, P. and Mariotti, M. (2007), Sequentially Rationalizable Choice. *The American Economic Review* 95, 1824-1839.

Matsushima, H. (2008a), Behavioral Aspects of Implementation Theory. *Economics Letters* 100, 161-164.

——— (2008b), Role of Honesty in Full Implementation. *Journal of Economic Theory* 139, 353-359.

Myerson, R. (1979), Incentive Compatibility and the Bargaining Problem. *Econometrica* 47, 61-74.

Repullo, R. (1985), Implementation in Dominant Strategies under Complete and Incomplete Information. *The Review of Economic Studies* 52, 223-229.

Rubinstein, A. and Salant, Y. (2008), (A ,f): Choice with Frames. *The Review of Economic Studies* 75, 1287-1296.

Saijo, T. (2005), Spiteful Behavior in Voluntary Contribution Mechanism Experiments. In *Handbook of Experimental Economics Results*, ed. by C. R. Plott and V. L. Smith. North-Holland.

Saran, R. (2011), Menu-Dependent Preferences and Revelation Principle. Forthcoming in *Journal of Economic Theory*, doi:10.1016/j.jet.2011.03.009.

Sen, A. (1977), Social Choice Theory : A Re-Examination. *Econometrica* 45, 53-89.

TURUN KAUPPAKORKEAKOULUN JULKAISUSARJASSA A OVAT VUODESTA 2010 LÄHTIEN ILMESTYNEET SEURAAVAT JULKAISUT

A-1:2010	Pikka-Maaria Laine Toimijuus strategiakäytännöissä: Diskurssi- ja käytäntöteoreettisia avauksia
A-2:2010	Juha V. Virtanen Johtajana sairaalassa. Johtajan toimintakenttä julkisessa erikoissairaalassa keskijohtoon ja ylimpään johtoon kuuluvien lääkäri- ja hoitajataustaisten johtajien näkökulmasta
A-3:2010	Markus Vinnari The past, present and future of eating meat in Finland
A-4:2010	Arja Lemmetyinen The coordination of cooperation in tourism business networks
A-5:2010	Antti Fredriksson Soft and hard information in bank-firm relationships - The effect of relationship intensity and reporting conservatism on loan pricing and relationship value
A-6:2010	Tomi Kortela The intertemporal choice of households under incomplete markets in general equilibrium
A-7:2010	Hongxiu Li E-service continuance: an insight into online travel services in China
A-8:2010	Harri Virolainen “Kai sitä ihminen on vaan semmoinen laumaeläin” – Virtuaalisen tiimin ilmapiiri
A-1:2011	Marjo Kumpula Vakuutusalan työn sisältö ja työntekijöiden ammatti-identiteetin muovautuvuus – Vakuutusvirkailijasta finanssialan myyjäksi?
A-2:2011	Helena Keinänen Algorithms for coalitional games
A-3:2011	Matti Mäntymäki Continuous use and purchasing behaviour in social virtual worlds
A-4:2011	Arto Kuuluvainen Dynamic capabilities in the international growth of small and medium-sized firms
A-5:2011	Ville Korpela Four essays on implementation theory

Kaikkia edellä mainittuja sekä muita Turun kauppakorkeakoulun julkaisusarjoissa ilmestyneitä julkaisuja voi tilata osoitteella:

KY-Dealing Oy
Rehtorinpellonkatu 3
20500 Turku
Puh. (02) 333 9422
E-mail: ky-dealing@tse.fi

All the publications can be ordered from

KY-Dealing Oy
Rehtorinpellonkatu 3
20500 Turku, Finland
Phone +358-2-333 9422
E-mail: ky-dealing@tse.fi