Rev. Econ. Design ♣. ♣–♣ 035

Review of Economic Design

Implementation with partial verification

Nirvikar Singh, Donald Wittman

Department of Economics, University of California, Santa Cruz, Santa Cruz, CA 95064, USA
(e-mail: wittman@cats.ucsc.edu)

Received: 1 August 1998 / Accepted: 5 September 2000

Abstract. This paper examines the implementability of social choice functions when only partial verification of private information is possible. Green and Laffont (1986) used this framework to derive a necessary and sufficient condition for the revelation principle to continue to hold with partial verification. We provide economically interesting characterizations of this condition, which suggest that it may be too restrictive. This leads us to consider implementation (not necessarily truthful) in general, when there is partial verification. We consider the case where compensatory transfers are allowed, giving the mechanism designer further leeway. We show how partial verification may allow efficient implementation of bilateral trade, where it would otherwise not be possible.

Key words: Implementation, partial verification, mechanism design, revelation principle

JEL classification: D8

1 Introduction

The implementation of desired policy often requires both monitoring and the proper design of incentives. But these may be incomplete: on the one hand, with a given set of preferences of the agents and goals of the principal, incentives may not be perfectly tuned; on the other hand, monitoring may not be able to distinguish beyond broad categories. In their seminal paper, Green and Laffont

We are grateful for comments from seminar audiences at CES Munich, UC Berkeley and the Naval Post Graduate School, Monterey. We are indebted to the two referees for helpful comments on the proper interpretation of the model and for helping us tighten up the mathematics and the verbal exposition. We received financial support from the Commitee on Research of the Academic Senate at UCSC.

(1986) examined the robustness of the revelation principle to the possibility of partial verification by the principal of the agent's message. Partial verification, there, referred to restrictions on the set of messages available to the agent–hence restrictions on the type or degree of misrepresentation. Green and Laffont derived a necessary and sufficient condition on the message space for the revelation principle to hold for every implementable social choice function.[1] This condition (which we state below) is called the nested range condition (NRC).

In this paper we expand on the Green-Laffont framework. After a brief review of their model and a discussion of their main results in Sect. 2, we provide in Sect. 3. some economically interesting characterizations of message correspondences that satisfy NRC. We show that economically plausible message spaces are likely to violate NRC. Therefore, when types are partially observable, mechanism designers should not limit their attention to truthful implementation. In Sect. 4, we provide a necessary and sufficient condition on the the message correspondence for implementability of a social choice function, when the sets of types and allocations are subsets of R, and where preferences are strictly monotone in the allocation. We discuss how this generalizes to R^n, provided that preferences satisfy a unanimity condition. We derive some more constructive special results for the one-dimensional case, showing when all social choice functions are truthfully implementable and when only the constant function is implementable. In Sect. 5, we examine the case where the allocation is still a real number, but transfers are possible, as well. This provides additional degrees of freedom to the mechanism designer. We first consider the case where there are a finite number of types and characterize situations where implementation of any social choice function is possible. We then consider the case where there is a continuum of types. We place some further restrictions on preferences (unanimity that more is better) and message correspondences, and provide an alternative method for constructing a mechanism that implements social choice functions when the usual conditions for implementability (without partial verification or restricted message spaces) are violated.

Finally, in Sect. 6, we consider bilateral trade. We show how restricted message spaces may enable market mechanisms to be efficient even when both parties to a transaction have private information. Thus, some kinds of partial verification overcome the inefficiency (identified by Myerson and Satterthwaite 1983) that occurs when there are no restrictions on the message spaces. Note that achieving this involves abandoning NRC and the revelation principle – implementation is not truthful. Section 7 provides a summary conclusion, including suggestions for further research.

2 The model

We first present the general model as specified by Green and Laffont, then introduce our particular assumptions. The mechanism design problem is thought of as

[1] For general statements of the revelation principle, see Dasgupta et al. (1979).

a principal-agent problem. The agent's utility function depends on a parameter, $\theta \in \Theta$, and on a decision or allocation $x \in X$, and is denoted $u(x, \theta)$. The agent observes θ and transmits a message regarding this to the principal, who then chooses an action x based on the message. The principal does not observe θ. The crucial feature of the model is that the message for an agent of type θ is restricted to a subset of Θ which may vary with θ.

The framework is formalized in the following sequence of definitions, which follows Green and Laffont closely.

D.1. $M : \Theta \to \Theta$ is a correspondence determining the admissible messages – the *message correspondence* – i.e. for each $\theta, M(\theta)$ is the set of messages the agent can send.[2] An element of $M(\theta)$ is denoted by m.
Green and Laffont also add the requirement that $\theta \in M(\theta)$, i.e., truthtelling is always feasible. We call this the G-L condition.

D.2. $g : \Theta \to X$ is a function chosen by the principal, specifying an allocation for each possible message – the *outcome function*.

D.3. A *mechanism* is a pair $(M(\cdot), g)$.[3]

D.4. Given a message correspondence M and an outcome function g, the agent's *response rule*, is a function $\phi_g : \Theta \to \Theta$ defined by $\phi_g \in arg \max_{m \in M(\theta)} u(g(m), \theta)$.[4]

D.5. A *social choice function*, $f : \Theta \to X$ specifies a decision x for each possible agent type θ.

Typically, one is interested in the nature of the allocations that can be made by the principal, given the asymmetry of information. Here the asymmetry is complicated by the message correspondence $M(\cdot)$. The analysis is further complicated by the fact that the set of possible decisions depends upon $M(\cdot)$. This is formalized in the next two definitions.

---

[2] Green and Laffont (p. 451) provide the following interpretation of $M(\theta)$. The principal observes a binary variable whose value is non-stochastically jointly determined by the true type θ and the message sent, m. The variable's value indicates whether or not $m \in M(\theta)$. Severe punishments for $m \notin M(\theta)$ ensure that no agent will ever send such a message. thsi interpretation differs from allowing the principal to see the entire set $M(\theta)$.

A trivial example occurs when a potential employer is requesting college transcripts. The applicant cannot claim more years of college than he/she actually undertook. If he/she does, it would mean immediate disqualification from the job. A more interesting example is found in the used car market. Since buyers have ready access to the *Blue Book of Used Car Prices*, a California seller of a 1994 Honda Accord with 60,000 miles on it cannot credibly claim that it is worth 12,000 dollars. But sellers with cars of above average condition in a category (poor, good, excellent) might be able to credibly claim that their car was in the next higher category.

We consider this interpretation further in the conclusion.

[3] We follow Green and Laffont in including M in the definition of a mechanism. However, as a referee has emphasized, M is exogenous in this framework, and not chosen by the mechanism designer.

[4] We follow Green and Laffont in making the response rule a function rather than a correspondence. If $\operatorname{argmax}_{m \in M(\theta)} u(g(m), \theta)$ is not a singleton, any selection can be made, resulting in an induced response rule.

D.6. f is *$M(\cdot)$-implementable* iff there exists an outcome function $g : \Theta \to X$ such that $g(\phi_g(\theta)) = f(\theta)$ for any $\theta \in \Theta$, where ϕ_g is an induced response rule.[5]

D.7. f is *truthfully $M(\cdot)$-implementable* iff there exists an outcome function $g^* : \Theta \to X$ such that, for any θ in Θ,
(i) $g^*(\phi_{g^*}(\theta)) = f(\theta)$ and (ii) $\phi_{g^*}(\theta) = \theta$, where ϕ_{g^*} is an induced response rule.

The second requirement is that of truthfulness. It is clear that if f is truthfully *$M(\cdot)$-implementable*, then it is *$M(\cdot)$-implementable*. In the standard principal-agent literature, $M(\theta) = \Theta$ for any θ in Θ, and the reverse implication also holds: any *M-implementable* social choice function is truthfully implementable. This result is known as the revelation principle. Green and Laffont extend the revelation principle to other $M(\cdot)$. They use the following condition.

The *Nested Range Condition (NRC)*. For any three distinct $\theta^1, \theta^2, \theta^3$ in Θ, if $\theta^2 \in M(\theta^1)$ and $\theta^3 \in M(\theta^2)$, then $\theta^3 \in M(\theta^1)$.[6]

Another way of stating NRC is that if $\theta^2 \in M(\theta^1)$, then $M(\theta^2) \subseteq M(\theta^1)$.

Green and Laffont's fundamental result is the following theorem.

Theorem. (Green and Laffont). *Given that the G-L condition holds ($\theta \in M(\theta)$ for all θ) :*

(1) If $M(\cdot)$ satisfies NRC, then, for any X and u: $X \times \Theta \to R$, the set of implementable social choice functions coincides with the set of truthfully implementable social choice functions.

(2) If $M(\cdot)$ violates NRC, then there exist X, u: $X \times \Theta \to R$, and an $M(\cdot)$- **implementable** *social choice function f such that f is not truthfully $M(\cdot)$-* **implementable.**

This theorem substantially extends the revelation principle to situations where the principal can partially verify the agent's message, so that the agent cannot plausibly transmit a message outside $M(\cdot)$. Green and Laffont consider several examples, some satisfying NRC, some not, in order to illustrate the scope of their result. These include cases where Θ is a closed interval of R. This is a common situation in economic applications. The parameter θ might be the agent's ability as measured by a test score, his income, or the probability that he will have an accident. Hence, while the G-L theorem applies to general spaces Θ, it is interesting and useful to more fully understand its implications in the case where $\Theta \subset R$. This is what we attempt in the next section of our paper. We shall also briefly discuss the case $\Theta \subset R^n$, and note the difficulties in extending the analysis to this case. However, the large literature dealing with situations where

[5] As emphasized by the referee, *$M(\cdot)$-implementation* is not the same as full implementation, in which $M(\theta) = \Theta$.

[6] In Sects. 2-4, we denote specific vectors by superscripts and particular elements of a given vector by subscripts. We deviate from this notation in Sect. 5.

the parameter θ is a real number suggests that this case alone is of sufficient interest.[7] We therefore begin with the following assumptions.

A.1. $\Theta = [0, 1] \subset R$.
Of course any closed interval would be sufficient – we have chosen [0, 1] for simplicity, since θ can be scaled appropriately.

A.2. For each $\theta, M(\theta)$ is a closed interval, $[\underline{m}(\theta), \bar{m}(\theta)] \subseteq [0, 1]$.
This is a reasonable assumption in many circumstances. For example, the principal may know that the agent's probability of having an accident is between 1/2 and 3/4. It is unlikely that the agent could transmit a probability of either between 1/2 and 5/8 or between 3/4 and 1, but not between 5/8 and 3/4.
We shall call $\underline{m}(\theta)$ and $\bar{m}(\theta)$ the *minimal* and *maximal* messages, respectively, for type θ. We sometimes assume that they satisfy the following property.

A.3. $\underline{m}(\theta)$ and $\bar{m}(\theta)$ are nondecreasing in θ.

This also seems reasonable. If the agent's true probability of having an accident is 5/8, and he can "get away with" any message between 1/2 and 3/4, then another person with a higher probability of having an accident should not be able to get away with an assessment of less than 1/2, and may, if he wishes, be able to declare that his true probability is greater than 3/4. As another example, a professor may under-report her true income by only reporting the readily verified wage income, but not reporting consulting income. But if these are highly correlated, the higher the true income the higher the reported income.

In the next section, we characterize NRC, given A.1.–A.3. We then discuss the implications of relaxing these assumptions. The characterization of NRC has a simple geometric representation, as well as an intuitive interpretation for economic examples that satisfy the assumptions. It also provides a basis for illustrating some results on implementability with partial verification, which are presented in Sect. 4.

3 Characterization of NRC

We first derive some properties of the minimal and maximal messages under NRC, but without A.3., the assumption that $\bar{m}(\theta)$ and $\underline{m}(\theta)$ are nondecreasing in θ.

Proposition 1. *Given NRC, the G-L condition, and A.1-A.2, then for any* $\theta^0 \in \Theta$,

(i) $\bar{m}(\theta^0) > \theta^0$ *implies* $\bar{m}(\cdot)$ *is nonincreasing to the right at* θ^0;
(ii) $\underline{m}(\theta^0) < \theta^0$ *implies* $\underline{m}(\cdot)$ *is nonincreasing from the left at* θ^0.

[7] The work here should not be confused with parametric environments where there is some information about a person's initial endowments separate from the message space. See Groves et al. (1987).

Proof.

(i) Suppose not. Then, for any interval $[\theta^0, \theta')$ there exists a δ such that $\theta^0 + \delta \in (\theta^0, \theta')$ and $\bar{m}(\theta^0) < \bar{m}(\theta^0 + \delta)$. This follows from the supposition that $\bar{m}(\cdot)$ is not nonincreasing to the right. Let $\theta' \leq \bar{m}(\theta^0)$. Then it is also the case that $\bar{m}(\theta^0) \geq \theta^0 + \delta$. Hence there exists a $\delta > 0$ such that $\theta^0 + \delta \leq \bar{m}(\theta^0) < \bar{m}(\theta^0 + \delta)$. The first inequality implies $\theta^0 + \delta \in M(\theta^0)$. The second inequality implies $\bar{m}(\theta^0 + \delta) \notin M(\theta^0)$. Hence NRC is violated, a contradiction.

(ii) Suppose not. Then, for any interval $(\theta', \theta^0]$, there exists a δ such that $\theta^0 - \delta \in (\theta', \theta^0)$ and $\underline{m}(\theta^0) > \underline{m}(\theta^0 - \delta)$. This follows from the supposition that $\underline{m}(\cdot)$ is not nonincreasing from the left at θ^0. Let $\theta' \geq \underline{m}(\theta^0)$. Then it is also the case that $\underline{m}(\theta^0) \leq \theta^0 - \delta$. Hence, there exists a δ such that $\theta^0 - \delta \geq \underline{m}(\theta^0) > \underline{m}(\theta^0 - \delta)$. The first inequality implies $\theta^0 - \delta \in M(\theta^0)$. The second inequality implies $\underline{m}(\theta^0 - \delta) \notin M(\theta^0)$. Hence NRC is violated, a contradiction. q.e.d.

Proposition 1 also tells us that if $\bar{m}(\theta)$ or $\underline{m}(\theta)$ is continuous at θ but does not coincide with θ, it must be nonincreasing at θ. Hence if it changes, it will decrease with θ. This is rather implausible. If we add A.3, then $\bar{m}(\theta)$ or $\underline{m}(\theta)$ must obviously be constant at such θ. However, we may prove the following result without A.3.

Proposition 2. *If NRC and A.1-A.2 are satisfied and $\underline{m}(\theta) < \theta < \bar{m}(\theta)$ for $\theta \in [\theta^0, \theta^1]$, then $\underline{m}(\theta)$ and $\bar{m}(\theta)$ are constant over the interval.*

Proof. By Proposition 1, $\underline{m}(\theta^1) \leq \underline{m}(\theta^0)$. Hence $\theta^0 \in M(\theta^1)$. Also by Proposition 1, $\bar{m}(\theta^1) \leq \bar{m}(\theta^0)$. Hence $\theta^1 \in M(\theta^0)$. Then from the restatement of NRC, $M(\theta^0) \subseteq M(\theta^1)$ and $M(\theta^1) \subseteq M(\theta^0)$; i.e. $M(\theta^1) = M(\theta^0)$; Since $\bar{m}(\cdot)$ and $\underline{m}(\cdot)$ are non-increasing, they must be constant over the interval. q.e.d.

Proposition 2 unfortunately tells us nothing when either no understatement ($\underline{m}(\theta) = \theta$) or, alternatively, no overstatement ($\bar{m}(\theta) = \theta$) is possible. It is then possible to construct many examples satisfying NRC that do not seem to have any interpretation in terms of an economic example. Furthermore, there is no simple geometric characterization of NRC, even in the one-dimensional case. However, the additional requirement A.3 allows a complete characterization that includes several economically interesting special cases.

For the statement of the general characterization result, we introduce the following notation. Let I_i be an element of a finite or countably infinite partition, P_1, of the interval $(0, 1]$ into intervals of the form $(\theta^i, \theta^{i+1}]$, i.e. $[0, 1] = \{0\} \cup (\cup_i I_i), i = 1, 2, \ldots$. Similarly, let J_j be an element of a finite or countably infinite partition, P_2, of the interval $[0, 1)$ into intervals of the form $[\theta^{*j+1}, \theta{*}^{j})$, i.e. $[0, 1] = (\cup_j J_j) \cup \{1\}, j = 1, 2, \ldots$.

The properties of $\bar{m}(\cdot)$ and $\underline{m}(\cdot)$ may be described on these partitions.

Proposition 3. *Suppose A.1–A.3 and the G-L condition holds. Then $M(\cdot)$ satisfies NRC if and only if there are partitions P_1 and P_2 such that* *(i)* $\bar{m}(0) = 0$ *or* θ^2, *and, for each* I_i, *either* $\bar{m}(\theta) = \theta$ *for any* $\theta \in I_i$ *or* $\bar{m}(\theta) = \theta^{i+1}$ *for any* $\theta \in I_i$.

(ii) $\underline{m}(1) = 1$ *or* θ^{*2} *and, for each* J_j, *either* $\underline{m}(\theta) = \theta$ *for any* $\theta \in J_j$ *or* $\underline{m}(\theta) = \theta^{*j+1}$ *for any* $\theta \in J_j$.

Proof. From Proposition 1 and A.3, it is clear that NRC implies that $\bar{m}(\cdot)$ and $\underline{m}(\cdot)$ must be step functions of the form described in the statement of the proposition, or that $M(\theta) = \{\theta\}$ over some intervals. We therefore show that any $\bar{m}(\cdot)$ and $\underline{m}(\cdot)$ of the above form insure that NRC is satisfied.

Consider say $\theta' \in \Theta$, and any $\theta'' \neq \theta'$ such that $\theta'' \in M(\theta')$. Either $\theta'' > \theta'$ or $\theta'' < \theta'$. Suppose $\theta'' > \theta'$. If $\theta' = 0$, then by assumption $\theta' < \theta'' \leq \theta^2 = \bar{m}(0)$. Hence $\theta'' \in (\theta^1, \theta^2]$ and $\bar{m}(\theta'') = \theta^2$. Furthermore, by A.3, $\underline{m}(0) \leq \underline{m}(\theta'')$. Hence, $M(\theta'') \subseteq M(\theta')$. If $\theta' > 0$, then $\theta' \in (\theta^i, \theta^{i+1}]$ for some i, and, by assumption, $\theta^i < \theta' < \theta'' \leq \theta^{i+1} = \bar{m}(\theta')$. Hence $\theta'' \in (\theta^i, \theta^{i+1}]$ and $\bar{m}(\theta'') = \theta^{i+1}$. Furthermore, by A.3, $\underline{m}(\theta') \leq \underline{m}(\theta'')$. Hence $M(\theta'') \subseteq M(\theta')$.

Now suppose $\theta'' < \theta'$. By similar reasoning, for $\theta' = 1$ or $\theta' \in J_j$, we can show that $M(\theta'') \subseteq M(\theta')$. But then for any $\theta' \in \Theta$ and $\theta'' \in M(\theta'), M(\theta'') \subseteq M(\theta')$. Hence NRC is satisfied. q.e.d.

The above proposition gives a complete characterization of NRC under some assumptions that we claim are of economic interest: one dimensional agent type and limits on messages that are decreasing in the agent's parameter. Some particular examples of this characterization illustrate plausible economic situations where there is partial verification. Rather than describe them algebraically, it is convenient to do so geometrically.

Example A. The most straightforward example of NRC (aside from $M(\theta) = \theta$ everywhere or $M(\theta) = \Theta$ everywhere) is the case where the principal's monitoring or verification technology allows him to divide agent types into distinct groups. This is illustrated in Fig. 1.A where the shaded area is $M(\cdot)$. It would fit well with an interpretation of imperfect observation or of aggregation of information.

Example B. An extension of the above would allow for some blurring of the boundaries between groups. Suppose types in the middle could safely pretend to be in the group on either end. Then we would have Figure 1.B. Note that by Proposition 1, the steps must occur at the truth line, $m = \theta$.

Example C. An obverse case would be where types in the middle can be identified, but extreme types cannot be so distinguished. This is illustrated in Fig. 1.C.

Example D As suggested by Green and Laffont, it may be possible for the agent to partially deceive the principal in one direction only, e.g. by overstating θ. This is shown in Fig. 1.D.

In this final example, whether the principal can perfectly discriminate or only screen into groups depends on the social choice function being implemented. Assuming u is increasing in x, if the social choice function is strictly decreasing in θ, each type of agent will state the lowest possible m, i.e. the truth, If f is strictly increasing in θ, each agent will give the highest possible message, and

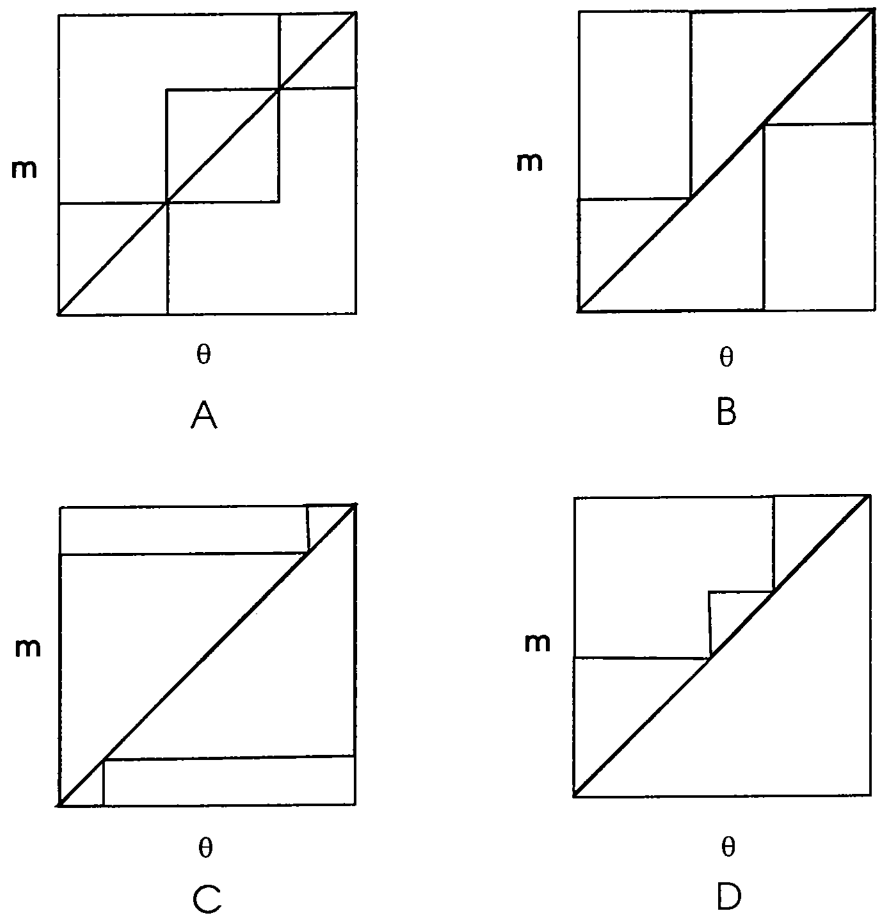

Fig. 1A–D. Examples of message spaces satisfying NRC

f will not be implementable. There is a parallel analysis, of course, for the case where the agents can only understate θ.

One can construct more complicated examples, but by Proposition 3, they will be extensions and combinations of the four examples above, possibly interspersed with intervals where $M(\theta) = \theta$, i.e., where perfect verification is possible. Without A.3 (the non-negative monotonicity assumption), it is possible to construct other sorts of examples satisfying NRC, but these seem implausible in any case. In the next section, we discuss implementability given cases of partial verification, such as illustrated above. Before that, we briefly discuss the extension of the above results to more general Θ.

The extension we consider is where $\Theta = \times_{i=1}^{n}[0, 1] \subset R^n$. This would describe situations where several numerical-valued characteristics are unobservable but reported by the agent. For example, this may be ability and income, or income in a sequence of years. If $M(\theta) = \times_{i=1}^{n} M_i(\theta_i)$, then the restriction on M_i depends only on θ_i, but not on any $\theta_j, j \neq i$. For example, the agent's claims about income (dimension i) may not depend for their plausibility on his true ability (dimension j). For this case, the analysis of one-dimensional agent characteristics extends forwardly.[8]

4 Implementability and truthful implementability

In the previous section we characterized NRC and showed that the requirements of NRC are unlikely to be satisfied in practice. Therefore designers will not be able to use the revelation principle under NRC as a guide in designing mechanisms. Instead they will have to look elsewhere. In this and the following section we provide alternative guides for mechanism design and show that when the message space varies across types, the focus on truthful implementability may be misplaced.

In this section we chiefly restrict attention to the case where the allocation space is one-dimensional and where utility is strictly increasing in the allocation. We will point out how this assumption can be relaxed for certain results. The first result of this section is a general one, without the structure of the message correspondence implied by A.1–A.3. It provides a necessary and sufficient condition for implementability of a social choice function. Further results are derived as consequences of Theorem 1, or with additional structure placed on the message correspondence. We also relate the possibility of implementation to the satisfaction of NRC.

Accordingly, we begin with

A.4. x is a real number and $u_x > 0$, where the subscript denotes a partial derivative.

We also introduce some notation before we state the first result.

Let $A_f(\theta') = \{\theta : f(\theta) < f(\theta')\}$. Hence $A_f(\theta')$ is the set of types who are worse off than θ', under the social choice f, since $u_x > 0$.

Also, let $K_f(\theta') = \cup_{\theta \in A_f(\theta')} M(\theta)$. Hence $K_f(\theta')$ is the set of all messages available to any of the types who are worse off than θ', given f.

Finally, let $M_f^*(\theta') = M(\theta') \cap K_f^c(\theta')$, where K_f^c is the complement of K_f. Thus $M_f^*(\theta')$ is the set of messages available to θ' but not to anyone worse off under f.

Theorem 1. *Given A.4, f is $M(\cdot)$-implementable if and only if $M_f^*(\theta')$ is nonempty for each $\theta' \in \Theta$.*

---

[8] A more general form of $M(\cdot)$ would be given by $M(\theta) = \times_{i=1}^{n} M_i(\theta)$. Here the feasible message for characteristic i depends on the whole vector θ. For example, an agent may not be able to plausibly claim a much lower income this year than last year, so that what he can get away with this year depends on his earnings in both years. Clearly, we cannot now analyze the problem one dimension at a time, and, while it is possible to construct various examples analogous to Examples A-D above, one cannot derive general results such as Proposition 3.

Proof.

(i) Sufficiency.

(a) We first show that if $f(\theta'') \neq f(\theta')$, then $M_f^*(\theta'') \cap M_f^*(\theta') = \emptyset$. Without loss of generality, suppose $f(\theta'') < f(\theta')$. But then $M(\theta'') \subseteq K_f(\theta')$. Since $M_f^*(\theta'') \subseteq M(\theta'')$ and $M_f^*(\theta') = M(\theta') \cap K_f^c(\theta'.)$ the result follows.

(b) Now for any θ', let the outcome function g be such that (1) $g(m) = f(\theta')$ for $m \in M_f^*(\theta')$ and (2) $g(m) < f(\theta')$ for $m \in M(\theta')|M_f^*(\theta')$, where A|B stands for the complement of B with respect to A. That is, under (b)(2), the designer allocates less than $f(\theta')$ for any message that is available both to θ' and another θ that is worse off under f.

From (a), this definition is consistent over all sets $M_f^*(\theta')$. Also, if $f(\theta'') < f(\theta')$, then $g(m) = f(\theta'')$ for $m \in M_f^*(\theta'')$ is consistent with the second part of the definition. Finally, if $f(\theta'') > f(\theta')$, then $M(\theta') \subseteq K_f(\theta'')$. So $M_f^*(\theta'') \cap M(\theta') = \emptyset$, and $g(m) = f(\theta'')$ for $m \in M_f^*(\theta'')$ is consistent with the second part of the definition.

Hence g is well-defined and implements f, since $u_x > 0$.

(ii) Necessity

Suppose f is $M(\cdot)$-implementable. Then it must be that $f(\theta') = g(\phi_g(\theta')) > g(\phi_g(\theta))$, for all $\theta \in A_f(\theta')$, where ϕ_g is the response rule, i.e.: $\phi_g(\theta) = arg \max_{m \in M(\theta)} u(g(m), \theta)$. That is, by definition of $A_f(\theta')$, the types in the set cannot attain $f(\theta')$.

But $u_x > 0$ implies $u(g(\phi_g(\theta')), \theta) > u(g(\phi_g(\theta)), \theta)$ for any $\theta \in A_f(\theta')$. Hence $\phi_g(\theta') \notin K_f(\theta')$, i.e., the message $\phi_g(\theta')$ is not available to those who are worse off than θ', given f. Hence $\phi_g(\theta') \in M_f^*(\theta')$; so this set must be nonempty. q.e.d.

Theorem 1 provides a substitute for the revelation principle (when $u_x > 0$). If and only if the mechanism designer can find a "unique" message for θ (with respect to the set of θ made worse off by the social choice function) can a non-constant social choice function be implemented. The theorem not only tells the designer where to look, but also tells the designer at the same time whether the social choice function is implementable in the first place. Of course, if the social choice function is implementable, the mechanism designer is able to infer the truth from the message.

It may be noted that the result of Theorem 1 does not require any of the structure on θ or $M(\theta)$ imposed in A.1–A.3. In fact, it does not even require truth-telling to be feasible, although we would expect it to be generally so! Finally, f need not be continuous or monotone for the result.

The assumption that $u_x > 0$ serves to give unanimity of rankings of outcomes by different types, and this is the critical requirement for Theorem 1. Hence the result will also hold for the more general case where $x \in R^n$ and $f : R^n \to R^n$, provided $u(x, \theta) = v(x)a(\theta) + b(\theta)$ where $a(\theta) > 0$. In this case, $u(x', \theta) - u(x'', \theta) = [v(x') - v(x'')]a(\theta)$, which has the same sign for all θ. The proof in this case involves working with the utilities rather than directly with f to construct the appropriate sets, e.g., $A_f(\theta') = \{\theta : u(f(\theta), \theta) < u(f(\theta'), \theta')\}$, etc.

Corollary 1.1. Given the conditions of Theorem 1, let $M^{\#}$ be a set of messages which are common to all θ. The deletion or addition of $M^{\#}$ will not have any effect on the implementability of any social choice function f.

Proof. The result follows from Theorem 1 since the deletion or addition of a common set of messages does not affect the structure of the sets $M_f^*(\theta)$.

The statement concerning deletion is equivalent to Green and Laffont's Theorem 3. The statement concerning addition holds despite Green and Laffont's counter-example following their Theorem 3 because of our assumption that u is monotonically increasing in x (or, more generally, that there is unanimity of ranking of allocations x for all θ). Green and Laffont's counter-example does not satisfy this condition.

In the remainder of this section, we assume that $\theta \in M(\theta)$, i.e., the truth is a feasible message.

Theorem 2. *Given A.1–A.2 and A.4 if $f(\theta)$ is a strictly monotonically increasing (decreasing), it is $M(\cdot)$-implementable if and only if $\bar{m}(\theta)(\underline{m}(\theta))$ is strictly increasing.*

Proof. We will prove the result for f increasing – the other case is very similar

(i) Sufficiency
This follows from Theorem 1, since $\bar{m}(\theta) \in M_f^*(\theta)$. In fact any nondecreasing f is implementable.
(ii) Necessity:
Assume g implements f. Consider any $\theta' \in \Theta$ and suppose $\phi_g(\theta') < \theta'$. Let $\theta'' = \phi_g(\theta')$. Since $f(\theta'') < f(\theta')$, by Theorem 1 it must be that $\phi_g(\theta') \notin K_f(\theta') \supseteq M(\theta'')$. But $\theta'' \in M(\theta'')$, a contradiction. Hence $\phi_g(\theta') \geq \theta'$ for each $\theta' \in \Theta$. By Theorem 1 again, $\phi_g(\theta') \notin M(\theta'')$ for any $\theta'' < \theta'$. Hence, using A.2, $\phi_g(\theta') > \bar{m}(\theta'')$. Hence $\bar{m}(\theta') > \bar{m}(\theta'')$ for any $\theta' > \theta''$. q.e.d.

Note that if $\bar{m}$ or $\underline{m}$ is strictly increasing, NRC may be violated.[9] The following corollaries are almost immediate consequences of Theorem 2.

[9] Under the tip rate determination plan, the U.S. Internal Revenue Service has restaurant owners estimate tips based on a fixed percentage of the price charged for meals. Waiters might still under-report tips, but the tip rate determination plan puts a lower bound on such under-reporting. To the degree that there is under-reporting by taxpayers in general, the tax rate is higher than it would be in the absence of under-reporting. So the effective tax rate is basically the same regardless of the truthfulness of the messages. Because the degree of under-reporting is not consistent across types, the tip rate determination plan only approximates the characterization of the message space in the proposition.

As an example closer to the academic world, professors may write honest letters of recommendation, but the letters themselves would be an upwardly biased estimate of the student's capabilities if the student chose the three professors who ranked the student the highest. But as long as the tendancy for upward bias is the same for all applicants, then graduate schools will admit the best students.

Corollary 2.1. *Given A.1–A.2 and A.4, if NRC holds and f is strictly increasing (decreasing) in θ, then f is $M(\cdot)$-implementable if and only if $\bar{m}(\theta) = \theta(\underline{m}(\theta) = \theta)$.*

Proof. From Theorem 2 and Proposition 1.

Corollary 2.2. *If $\bar{m}(\theta)$ is continuous and the conditions of Theorem 2 hold, f is uniquely $M(\cdot)$-implementable, where unique means only one message is possible for each θ.*

Proof. From the necessary conditions in Theorem 1.

Corollary 2.3. *If $\bar{m}(\theta) = \theta$, the conditions of Theorem 2 hold, and f is a non-decreasing function of θ, then f is uniquely and truthfully $M(\cdot)$-implementable.*

Proof. From Corollary 2.2.

The next result identifies the extreme case of partial verifiability that allows all social choice functions to be implementable. This turns out to be complete verifiability.

Theorem 3. *Given A.1–4, $M(\theta) = \{\theta\}$ is the only message correspondence which allows all social choice functions f to be $M(\cdot)$-implementable.*

Proof. If $M(\theta) = \{\theta\}$ obviously any f is $M(\cdot)$-*implementable*. We show that for any other $M(\cdot)$ there exists an f that is not $M(\cdot)$-*implementable*. If $M(\theta) \neq \{\theta\}$ for some $\theta \in \Theta$, there exists a $\theta' \in \Theta$ such that $\theta'' \in M(\theta'), \theta'' \in \Theta$ and $\theta'' \neq \theta'$.

Let $f(\theta) = \begin{cases} k & \text{for } \theta \in \Theta | \{\theta''\} \\ k + \epsilon & \text{for } \theta = \theta'' \end{cases}$

where $\epsilon > 0$. If f is $M(\cdot)$-*implementable* by g, it must be that $\phi_g(\theta'') \neq \theta''$, since $\theta'' \in M(\theta')$ and $f(\theta') = k$. Let $\theta''' = \phi_g(\theta'')$. Since $\theta''' \in M(\theta'''), \phi_g(\theta''') = \theta'''$. Hence $f(\theta''') = k + \epsilon$, a contradiction. Therefore the proposed f is not $M(\cdot)$-*implementable.* q.e.d.

We end this section with a discussion of the following question: what is the class of social choice functions that is implementable for all message correspondences that satisfy NRC? If the unanimity condition on preferences, $u(x, \theta) = v(x)a(\theta) + b(\theta)$ is satisfied, it is easy to see that the only such social choice function is $f(\theta) = c$, a vector of constants. This follows since $M(\theta) = \Theta$ satisfies NRC, and in this case all types can give the same message: hence they must receive the same allocation.

What is interesting is that if the preferences are not as above, other social choice functions are implementable. A simple example is where θ is a vector and $u(x, \theta) = -\sum_{i=1}^{n}(x_i - \theta_i)^2$. Then $f(\theta) = \theta$ is $M(\cdot)$-*implementable* as long as the truth is feasible for all types. Note that the dimensions of the allocation and message spaces are the same in this example.

Expanding on this idea, if $M(\theta) = \Theta$ for all θ and $f(\theta) \neq f(\theta')$ for $\theta \neq \theta'$, then $f(\theta)$ is truthfully implementable only if there is no unanimity at all.

5 Implementation with compensatory transfers

In this section we consider the case where compensatory transfers are possible. These enter utility in a linear separable fashion, but do not affect the allocation function, $f(\theta)$. Hence, in this section, utility is quasilinear of the form $u(x, \theta)+t$, $x, t \in R$, where $x = f(\theta)$ is the allocation, t is the compensatory transfer, and u is bounded. Since t can be tailored for individual θ, compensatory transfers allow additional degrees of freedom to the designer of the mechanism. In this section and the next, a mechanism is a triple $(M(\cdot), g, t)$.

As in the previous section, the relationship between the message correspondences plays a key role. Once again, mechanism design relies on messages that are unique with regard to *particular* subspaces. We will now explicitly describe the mechanism through an iterative, constructive process.

(5.1) For any $\theta \in \Theta$, let $M_1(\theta) = \{m : m \in M(\theta) and m \notin \cup_{\theta' \neq \theta} M(\theta')\}$.
If $M_1(\theta)$ is nonempty, then type θ has a message available that is not available to any other type, and $M_1(\theta)$ is the set of such messages.

(5.2) Let $M_1 = \cup_\theta M_1(\theta)$. This is the set of messages uniquely available to some type.

(5.3) Finally, let $S_1 = \{\theta : M_1(\theta) \neq \varnothing\}$. This is the set of types with unique messages.
Suppose that S_1 is nonempty, i.e. at least one $M_1(\theta)$ is nonempty.

Consider $\Theta_2 = \Theta | S_1$ (with $\Theta_1 = \Theta$). This is the set of types without unique messages.

For any $\theta \in \Theta_2$, let $M_2(\theta) = \{m : m \in M(\theta)$ and $m \notin \cup_{\Theta_2 | \{\theta\}} M(\theta')\}$. Again, these are messages uniquely available to type θ, but with respect to the restricted set Θ_2.

Let $M_2 = \cup_{\Theta_2} M_2(\theta)$ and let $S_2 = \{\theta : M_2(\theta) \neq \varnothing\}$.

If S_2 is nonempty, we continue recursively, as follows:

$\Theta_i = \Theta_{i-1} | S_{i-1}$.

For any $\theta \in \Theta_i$, let $M_i(\theta) = \{m : m \in M(\theta)$ and $m \notin \cup_{\Theta_i | \{\theta\}} M(\theta')\}$.

$M_i = \cup_{\Theta_i} M_i(\theta)$ and let $S_i = \{\theta : M_i(\theta) \neq \varnothing\}$.

Suppose there exists an $i = n - 1$ such that $\Theta_n = \Theta_n | S_n$ is a singleton, say $\{\theta''\}$. Then $M_n(\theta'') = \{m : m \in M(\theta'')$ and $m \notin \cup_{\Theta_n | \{\theta''\}} M(\theta')\}$ is nonempty, since $M(\theta'')$ is nonempty, and $\Theta_n | \{\theta''\} = \varnothing$, so the second condition is trivially satisfied.

Hence $M_n = M_n(\theta'')$ and $S_n = \{\theta''\}$.

Hence, without loss of generality, we can consider the case where there exists an n such that $\Theta = \cup_{i=1}^n S_i$.

In words, there is a finite partition of types such that each member S_j of the partition consists of types who have a message or messages not available to any type in S_i, $i > j$. Note also that $M_i \cap M_j = \varnothing$ for $i \neq j$.

Our first result regards implementation. Having partitioned Θ into sets of the above form, the mechanism designer can offer a significantly larger transfer to types in S_i than to types in S_{i+1}. This transfer will induce members of S_i to choose messages in M_i even if they are less satisfied with their allocation through the function, f. On the other hand, members of sets S_{i+1}, S_{i+2}, etc., do not have the messages in M_i available to them. Hence the key to implementability based solely on message correspondences is to look for messages that are unique within certain well defined subsets. We will now proceed with the theorem and formal proof.

Theorem 4. *If there is a finite partition of Θ such as described above, then every $f(\theta)$ is $M(\cdot)$-implementable.*

Proof. Consider S_n and any $\theta \in S_n$, say θ_n. For $f(\theta_n)$, choose some $m \in M_n(\theta_n)$, say m_n.
Let $g(m_n) = f(\theta_n)$.

All other messages in $M_n(\theta_n)$ can be assigned $-K$ where $-K < \inf_{\Theta} f(\theta)$. None of those messages will be relevant for anyone in $S_i, i < n$. This can be done for all $\theta \in S_n$. Since, by construction, m_n is not available to any other $\theta \in S_n$, $f(\theta_n)$ can be implemented within S_n without considering transfers. Therefore let the transfer to all θ in S_n be zero.

Now consider S_{n-1} and any $\theta \in S_{n-1}$, say θ_{n-1}. For $f(\theta_{n-1})$, choose some $m \in M_n(\theta_{n-1})$, say m_{n-1}. Let $g(m_{n-1}) = f(\theta_{n-1})$.

Now, within S_{n-1}, we have perfect screening without transfers, but it is possible that there is a θ_{n-1} and an $m \in M(\theta_{n-1})$ such that $m \in M_n(\theta_n)$ for some θ_n, i.e. some type in S_{n-1} may be able to profitably pretend to be a type in S_n. We need to use transfers to rule this out.

Then, for $f(\theta)$ to be implemented for θ_{n-1}, we require that $u(g(m_{n-1}), \theta_{n-1}) + t_{n-1} \geq u(g(m_n), \theta_{n-1})$ for all such m_n, or
$t_{n-1} \geq \sup_{m_n \in M(\theta_{n-1}) \cap M_n} u(g(m_n), \theta_{n-1}) - u(g(m_{n-1}), \theta_{n-1})$.

Furthermore, this should be true for all $\theta_{n-1} \in S_{n-1}$.
Hence we require

$$t_{n-1} \geq \sup_{\theta_{n-1} \in S_{n-1}} \left[\sup_{m_n \in M(\theta_{n-1}) \cap M_n} u(g(m_n), \theta_{n-1}) - u(g(m_{n-1}), \theta_{n-1}) \right].$$

The last condition ensures that no type in S_{n-1} has an incentive to pretend to be a type in S_n. As long as the suprema are finite (which holds if u is bounded, which we have assumed), t_{n-1} is a finite real number.

The transfers can be defined recursively now, as follows:

$t_i - 1 \geq \sup_{\theta_{i-1} \in S_{i-1}} \left[\sup_{m_j \in M(\theta_{i-1}) \cap M_j} u(g(m_j), \theta_{i-1}) - u(g(m_{i-1}), \theta_{i-1}) \right] + t_j$ for every $j \geq i$.

Since each supremum is finite, we can define a sequence of finite transfers, $\{t_1, ..., t_n\}$, which together with $g(m_i) = f(\theta_i)$ for some $m_i \in M_i(\theta_i), m$ implements for all $\theta \in \Theta$. q.e.d.

The theorem provides conditions under which *any* allocation function, f, is implementable. This is possible because (a) restrictions on message spaces and

(b) the possibility of compensatory transfers allow perfect screening regardless of preferences. Restrictions on relative valuations of the good x, such as the single crossing property, are not needed. The result can also be related to Theorem 1, where we showed that if $u_x > 0$ and those receiving more x had messages not available to those receiving less x, all social choice functions satisfying this condition were implementable. Here u_x need not be positive – compensatory transfers allow the mechanism designer to swamp differences based on different allocations of x. More explicitly, if the set of message spaces satisfy certain criteria, the mechanism designer can with the aid of transfers create a ranking of utilities such that those who receive higher utility will have messages available to them which are not available to those who receive lower utility. The sign of u_x does not matter.

We have two corollaries of Theorem 4.

Corollary 4.1. *If the conditions of Theorem 4 and the G-L condition hold and* $S_i \subseteq M_i$ *for* $i = 1, ..., n$, *then all* $f(\theta)$ *together with all* $t(\theta)$ *satisfy the construction in the theorem and are truthfully implementable.*

Proof. First consider S_1. All θ in S_1 have unique messages. By the definition of $M(\cdot)$ and the G-L condition, such messages belong to some $\theta \in \Theta$ and $\theta \in M(\theta)$. Therefore each $\theta \in S_1$ has one and only one message, m = θ. Suppose to the contrary. Then θ would have a message θ', but by the G-L condition, this message θ' should be available to θ'. Hence, implementability requires truthful implementability in this case.

By the above construction, no $\theta \in \Theta|S_1$ has a message, m, such that $m \in S_1$. Therefore the G-L condition holds for $\Theta|S_1$, and we can iterate the process conducted in the previous paragraph.

Corollary 4.2. *If* Θ *is a finite subset of the reals, and either*

(a) $\max_m\{m : m \in M(\theta)\} = \theta$ *for all* θ, *or*
(b) $\min_m\{m : m \in M(\theta)\} = \theta$ *for all* θ,

then all allocation functions, $f(\theta)$, *are truthfully implementable.*

Proof. The construction of the partition is obvious, starting with the highest type in case (a) or with the lowest type in case (b).

Remark 1. Here we do not require $M(\theta)$ to be an interval, as we did in Sect. 4.

Corollary 4.2 can be compared to Corollary 2.3, where non-decreasing f are truthfully implementable if $\bar{m}(\theta) = \theta$, and to Theorem 3, which states that all allocation functions are implementable only if $M(\theta) = \{\theta\}$ for all θ. In Theorem 4 and its corollaries, the possibility of transfers allows for considerably more freedom in implementation, both in terms of social choice functions and allowable message spaces. However, corollary 4.2 is restricted to a finite set of types.

It is also noteworthy that the conditions of Theorem 4 do not imply NRC, nor are they implied by NRC. For example, if $\Theta = \{\theta_1, \theta_2, \theta_3\}, M(\theta_1) = \{\theta_1, \theta_2\}$,

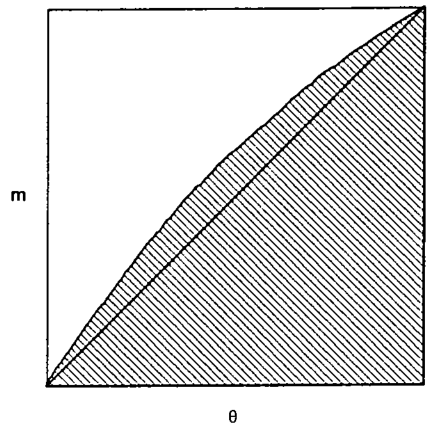

Fig. 2. Type $\theta = 1$ has a unique truthful message. However subsequent type's "unique" message (unique with respect to the running subset of types) is not the truth

$M(\theta_2) = \{\theta_2, \theta_3\}$, $M(\theta_3) = \{\theta_3\}$, then the conditions of Theorem 4 are satisfied, but NRC is violated. Conversely, if $M(\theta_1) = M(\theta_2) = \{\theta_1, \theta_2, \theta_3\}$, and $M(\theta_3) = \{\theta_3\}$, NRC is satisfied, but the conditions of Theorem 4 are violated. This is also the case for Corollary 4.2.

Corollary 4.1 implies that if the G-L condition holds and there is a finite set of types, then $M_i(\theta) = \theta$. Hence each type's set of unique messages can only be the truth. That this is not true for Θ infinite can be seen from Fig. 2. In this case $\theta = 1$ has a unique message which is the truth, but each subsequent type's "unique" message (unique with respect to the remaining subset of types) is not the truth.

Remark 2. The partition creates a natural ordering of the sets S_i, from $i = 1$ to n. When this ordering (or transitivity) is violated, a circularity is possible. Here, as in other areas of social choice, circularity undermines the relationships.

There are two problems with extending Theorem 4 to the case of an infinite partition. If the partition is countable, it may be that the transfers required become infinite and furthermore, there is no last element S_n, to begin the constructive proof. If the partition is uncountable, as in Figure 2, we cannot use the sequential approach at all.

Given the difficulty of extending the above approach to the infinite case, we develop an alternative method that places more structure on the problem and also yields more insight. The solution is, in some respects, quite similar to the standard approach without restrictions on message spaces, as we explain below.

Accordingly, let $\Theta = [0, 1]$ as in Sect. 2 and 3. Assume that $u_x > 0$, so the good to be allocated is always desirable. The usual assumption, typically termed the single crossing property, is that $u_{x\theta}$ does not change sign.

The necessary local condition for truthtelling (see, for example, Guesnerie and Laffont, 1984) is

(5.4) $u_{x\theta}(f(\theta), \theta) \cdot f'(\theta) \geq 0$