Game Theory with Applications to Finance and Marketing, II

Basic Ideas in Implementation Theory

Instructor: Chyi-Mei Chen
(Tel) 3366-1086
(Email) cchen@ccms.ntu.edu.tw
(Website) http://www.fin.ntu.edu.tw/∼cchen/

1. This note consists of five parts, following a description of notation and definitions. Part 1 introduces dominance strategy implementation, Part 2 Nash implementation, Part 3 subgame-perfect Nash implementation, Part 4 Bayesian implementation, and Part 5 briefly discusses the durability of a Bayesian mechanism.

2. **Notation and Definitions**.

3. There are I agents, $i = 1, 2, \cdots, I$. Let I denote the set $\{1, 2, \cdots, I\}$ as well. Let A denote the set of feasible social choices (or social states or outcomes). The profile of the I agents' preferences on A is state-dependent, where in state $\theta \in \Theta$, agent i has a preference ordering $R_i(\theta)$ on A (which defines the strict preference $P_i(\theta)$ and indifference $I_i(\theta)$). Agent i may each possess some information about θ; let his information be $\{\theta_i, R_i(\theta)\}$, and $\theta = (\theta_i)_{i\in I}$; i.e. there is no aggregate uncertainty facing the I agents in the sense that they would all learn θ if they could share information with one another. There are various possibilities regarding the relationships between the θ_i's and θ, and regarding the structure of Θ.

   - Agents may have *complete information*, if they all observe θ. A sufficient condition is that for all $\theta \in \Theta$, $\forall i, j \in I$, $\theta_i = \theta_j$. In general, if we assume that the structure of Θ and the set of prior beliefs of the I agents about elements in Θ are the agents' common knowledge, then we are in a Bayesian framework where agent i given his information $\{\theta_i, R_i(\theta)\}$ and his own prior $h_i(\theta)$ will be able to form a posterior $h_i(\theta_{-i}|\theta_i, R_i(\theta))$ for θ_{-i} (which is the $(I-1)$-vector obtained from deleting θ_i from θ). Here note that heterogeneity in prior beliefs is allowed.

- Agent i's preference may be independent of θ_{-i}, and he observes only θ_i. In this case, since θ_i completely determines $R_i(\theta_i)$, agent i's inference about θ_{-i} based on θ_i alone is the same as that based on both θ_i and $R_i(\theta_i)$. Note that if instead R_i depends on θ and agent i observes both θ_i and $R_i(\theta)$, then in general neither θ_i nor $R_i(\theta)$ is redundant when agent i tries to make inference about θ_{-i}.
- The set Θ may be either *unrestricted* (or call it the *universal* domain), or restricted in various ways. In particular, if $\Theta = \Pi_{i\in I}\Theta_i$ then we say preferences in Θ have independent domains. Note that statistic independence of θ_i's implies immediately independence in domains, but the converse is generally not true. If Θ is the universal domain, then certainly preferences have independent domains.

4. As Dasgupta, Hammond, and Maskin (1979) point out,[1] when moving from state θ to state θ', an outcome feasible in state θ may cease to be feasible in θ'. This implies that in general the set of feasible outcomes is state-dependent, say $A(\theta) \subset A$ in state θ. We shall however concentrate on the case where A is state-independent.

5. A social choice rule (SCR) or a social choice correspondence (SCC) is defined as a function $f : \Theta \to 2^A - \{\emptyset\}$; an SCC is a correspondence mapping elements of Θ into non-empty subsets of A.

6. There is a central planner endowed with an SCC, and his problem is to design a game form that implements f in a sense to be made precise below. In the sequel, unless stated explicitly, it will be assumed that (i) (independent domains) $\Theta = \Pi_{i\in I}\Theta_i$; and (ii) $R_i(\cdot)$ depends only on $\theta_i, \ \forall i \in I$.

7. A (static) <u>game form</u> or <u>mechanism</u> is a pair (g, S), where $S = S^1 \times S^2 \times \cdots \times S^I$ specifies the set of feasible (pure) strategy profiles and

[1] Dasgupta, D., P. Hammond, and E. Maskin, 1979, The implementation of social choice rules: Some general results on incentive compatibility, *Review of Economic Studies*, 185-216.

$g : S \to A$ maps each strategy profile into a social outcome in A. Note that the game form (g, S) differs from a normal-form game in that the agents' preferences on A are left unspecified. Given any $\theta \in \Theta$, (g, S, θ) becomes a well-defined normal-form game.

8. Two things are worth noting at this point. First, agents will then be playing *the same game form* across all preference states. Second, the reason that a game form instead of a game is designed at this point is because the central planner does not know θ; he only knows that θ is some element of Θ. By choosing S and g, the central planner realizes that different equilibrium outcomes may arise from this same game form when preference states differ, and he would like to ensure that with (g, S), in any state $\theta \in \Theta$, any equilibrium that the agents may reach, denoted by $s^*(\theta)$, is such that $g(s^*(\theta)) \in f(\theta)$. What has been imprecise is the solution concept used to "define" the equilibrium of the game (g, S, θ). In this note, we shall consider Nash equilibrium, SPNE, BE, and dominance equilibrium, and depending on the chosen equilibrium concept, the central planner's problem may be then referred to as one of Nash implementation, subgame-perfect Nash implementation, Bayesian implementation, or dominant strategy implementation.

9. Fix a solution concept, let $s^*(\theta)$ be one equilibrium of the game (g, S, θ). Let $E_g(\theta)$ be the set of all such equilibria in state θ. Then, define $g(E_g(\theta)) \equiv \{g(s^*) : \; s^* \in E_g(\theta)\}$.

10. A game form (g, S) is said to *implement* an SCC f, if for all $\theta \in \Theta$, (i) $E_g(\theta)$ is non-empty; and (ii) $g(E_g(\theta)) \subset f(\theta)$. The game form (g, S) is said to *fully implement* an SCC f, if for all $\theta \in \Theta$, (i) $E_g(\theta)$ is non-empty; and (ii) $g(E_g(\theta)) = f(\theta)$.

11. A game form (g, S) is said to be *direct*, if for all $i \in I$, $S^i = \Theta_i$. A direct game form (g, Θ) is a *direct revelation mechanism*, if given the specified solution concept, $\theta \in E_g(\theta)$ for all $\theta \in \Theta$. That is, turth-telling forms an equilibrium of the game (g, Θ, θ) for all $\theta \in \Theta$.

12. Fix a solution concept. Given game form (g, S), suppose $E_g(\cdot) : \Theta \to S$ is a well-defined nonempty *equilibrium correspondence*. An *equilibrium selection* of (g, S), denoted by $s^* : \Theta \to S$, is a mapping such that for all $\theta \in \Theta$, $s^*(\theta) \in E_g(\theta)$.

13. Given an indirect game form (g, S), a direct game form (h, Θ) is one of its *equivalent* direct mechanisms if (i) for all $\theta \in \Theta$, $\theta \in E_h(\theta)$; and (ii) for some equilibrium selection $s^*(\cdot)$, $h(\theta) = g(s^*(\theta))$, $\forall \theta \in \Theta$. For example, suppose that $\Theta = \{\theta, \theta'\}$, $\#E_g(\theta) = 2$, $\#E_g(\theta') = 3$, then there are six possible equilibrium selections, and each equilibrium selection defines one distinct equivalent direct game form for (g, S).

14. A direct mechanism (h, Θ) is said to implement f *truthfully* if for all $\theta \in \Theta$, (i) $\theta \in E_h(\theta)$ (given the adopted solution concept); and (ii) $h(\theta) \in f(\theta)$.

    By definition, if (g, S) implements f and (h, Θ) an equivalent direct mechanism of (g, S), then (h, Θ) truthfully implements f. Truthful implementation is a necessary (but insufficient) condition for implementation.

15. Note that if (g, S) implements an SCC f and (h, Θ) one of its equivalent direct game forms (so that (h, Θ) truthfully implements f), then (h, Θ) need not implement f; see section 22.

    Even if a direct mechanism (h, Θ) implements an SCF f *truthfully*, so that $h(\theta) = f(\theta)$, $\forall \theta \in \Theta$ and hence $h(\cdot)$ and $f(\cdot)$ are the same function (h, Θ) may still fail to (fully) implement f. In the latter case, we need, given any $\theta \in \Theta$, the set of equilibrium points of the game (h, Θ, θ) coincide with $f(\theta)$.

16. A direct game form (h, Θ) is a *straightforward*, if for all $i \in I$, all $\theta_i, \eta_i \in \Theta_i$, and all $\theta_{-i} \in \Theta_{-i}$,

    $$h(\theta_i, \theta_{-i}) R_i(\theta_i) h(\eta_i, \theta_{-i});$$

    that is, if truth-telling forms a *dominance equilibrium* of the game (h, Θ, θ) for all $\theta \in \Theta$.

17. An SCC $f : \Theta \to 2^A$ is *monotonic* if and only if: for all $\theta, \phi \in \Theta$ and for all $a \in A$, if $a \in f(\theta)$ and if for all $i \in I$, for all $b \in A$, $aR_i(\theta)b \Rightarrow aR_i(\phi)b$, then $a \in f(\phi)$.

18. Part 1. **Dominant Strategy Implementation.**

19. Given game form (g, S), s_i^* is a dominant strategy for i given θ_i if: $\forall s_i \in S^i$ and $\forall s_{-i} \in S^{-i}$,

$$g(s_i^*, s_{-i}) R_i(\theta_i) g(s_i, s_{-i}).$$

The game form (g, S) is a *dominant strategy mechanism* for Θ if, for all $i \in I$ and for all $\theta_i \in \Theta_i$, there exists a dominant strategy for agent i. (Green and Laffont (1979) called such a mechanism a *decisive* mechanism.)

20. There of course may be more than one dominant strategy for i in (g, S, θ). However, for a dominant strategy mechanism, there exists a *dominant strategy selection* $s^* : \Theta \to S$ (with $s^*(\theta) \equiv (s_i^*(\theta_i))_{i \in I}$) such that for each $i \in I$ and each $\theta_i \in \Theta_i$, $s_i^*(\theta_i)$ is a dominant strategy for i given θ_i.

21. (**Revelation Principle**) Let (g, S) be a dominant strategy mechanism. For each dominant strategy selection $s^* : \Theta \to S$ there exists a straightforward mechanism which is equivalent to (g, S). Hence, if (g, S) implements an SCC f in dominant strategy, then there exists a direct mechanism implementing f truthfully in dominant strategy. In short, an SCC which is implementable in dominant strategies is truthfully implementable.

Proof Pick one dominant strategy selection s^* and define $h : \Theta \to A$ by $h(\theta) = g(s^*(\theta))$ for all $\theta \in \Theta$. For any fixed $\eta_{-i} \in \Theta_{-i}$ and every $\eta_i \in \Theta_i$, we have by definition

$$s^*(\eta_i, \eta_{-i}) = (s_i^*(\eta_i), s_{-i}^*(\eta_{-i})).$$

So, for every θ_i, since $s_i^*(\theta_i)$ is a dominant strategy for i given θ_i, we have

$$g(s_i^*(\theta_i), s_{-i}^*(\eta_{-i})) R_i(\theta_i) g(s_i^*(\eta_i), s_{-i}^*(\eta_{-i})),$$

or equivalently,

$$h(\theta_i, \eta_{-i}) R_i(\theta_i) h(\eta_i, \eta_{-i}).$$

Thus truth-telling is a dominant strategy for (h, Θ); i.e. the latter is a straightforward mechanism.

22. The notions of truthful implementation is very weak. In moving from the indirect mechanism (g, S) to direct ones, one may introduce dominant strategies which are not truthful. More troubling, even if (g, S) does implement f, the equivalent direct mechanism (h, Θ) may fail to do so, as the following example shows.

Example D. Suppose that $A = \{a, b, c, d, e, p, q, r\}$, $\Theta_1 = \{R_1, R_1'\}$, $\Theta_2 = \{R_2, R_2'\}$. These preferences are summarized in the following table, where an upper row of decisions is preferred to a lower row.

R_1	R_1'	R_2	R_2'
q	c,b,p	r	d
a,c,e	a,d,e	d,a,e	b,c
d,b,p	q,r	b,c,p	a
r		q	e,p,q,r

Consider the SCC defined by $f(R_1, R_2) = \{a, e\}$, $f(R_1', R_2) = \{c, p\}$, $f(R_1, R_2') = \{d, b\}$, $f(R_1', R_2') = \{b\}$. Note that f is Peratian (i.e., $f(\theta)$ never contains elements of A that are Pareto dominated in state θ) and monotonic.

The following game form (g, S) implements f in dominant strategy:

a	d	e
c	b	p
a	b	e

In this table, S^1 consists of the three rows, and S^2 the three columns. Given R_1, agent 1 has the 1st and the 3rd rows as dominant strategies, and the 2nd row when given R_1'. Given R_2, agent 2 has the 1st and the 3rd columns as dominant strategies, and the 2nd column as the dominant strategy when given R_2'. Thus, for example, the game $(g, S, (R_1, R_2))$ has 4 dominance equilibria, and the game $(g, S, (R_1', R_2'))$ has 1 dominance equilibrium. There are in total 16 possible equilibrium selections, which define correspondingly 16 equivalent direct game forms.

We shall consider a special equivalent direct game form. Consider the equivalent direct mechanism (h, Θ) obtained from (g, S) by associating rows 1 and 2 of (g, S) with respectively R_1 and R'_1 and columns 1 and 2 of (g, S) with respectively R_2 and R'_2.

	R_2	R'_2
R_1	a	d
R'_1	c	b

This direct game form (h, Θ) does not implement f: in state (R_1, R_2), the game has an unwanted dominance equilibrium where agents play respectively the 2nd row and the 2nd column, resulting in b, but b is not contained in $f(R_1, R_2)$!

23. The important message that the preceding example delivers is that to implement an SCC we cannot count on direct game forms; creatively designed indirect game forms are needed. We shall demonstrate that this is indeed the case even when we adopt the other equilibrium concepts.

24. A closer inspection of the preceding example reveals that the unwanted equilibrium may arise because an agent has a non-singleton indifference set. By restricting preferences to be *strong* (indifference curves are all singletons), we may be able to get rid of the unwanted equilibrium problem. This is indeed true; see Theorem D2 below.

25. **Theorem D1.** Suppose that $R(\Theta)$ contains only strong orderings. Then, if an SCC f is fully implementable in dominant strategies, then f is an SCF.

Proof For $\theta \in \Theta$, consider $a, b \in f(\theta)$. We shall show that $a = b$. By assumption, some (g, S) fully implements f in dominant strategy, and so there must exist $s, s' \in S$ such that $g(s) = a$ and $g(s') = b$. Since for agent 1, given θ_1, s_1, s'_1 are both dominant strategies in state θ and since agent 1 never feels indifferent about any two distinct elements in A, we must have

$$a = g(s) = g(s_1, s_{-1}) = g(s'_1, s_{-1}).$$

Similarly, from agent 2's perspective, we have

$$g(s'_1, s_{-1}) = g(s'_1, s_2, s_{-\{1,2\}}) = g(s'_1, s'_2, s_{-\{1,2\}}).$$

Continuing iteratively, we have

$$a = g(s) = g(s') = b.$$

26. **Theorem D2.** Suppose that $R(\Theta)$ contains only strong orderings. If f is truthfully implemented in dominant strategies by the direct mechanism (h, Θ), then (h, Θ) implements f.

    Proof Define a correspondence $F : \Theta \to A$ by $F(\theta) \equiv h(E_h(\theta))$, where the solution concept is dominance equilibrium. By construction, (h, Θ) fully implements the SCC F, and thus by Theorem D1, F must be an SCF! Since $h(\theta) \in f(\theta)$ by truthful implementation, we conclude that $F(\theta) \in f(\theta)$, and hence by definition (h, Θ) implements f.

27. **Theorem D3.** Suppose that $R(\Theta) = \mathcal{P}^I$ and $\#(A) \geq 3$ is finite. Suppose that an SCF is Paretian and can be implemented in dominant strategies. Then, f is dictatorial.

28. Without requiring f to be an SCF, there can exist non-dictatorial Paretian SCC that can be implemented in dominance equilibrium. The SCC presented in Example D is monotonic and Paretian, and is fully implemented by the game form (g, S). However, recall that preferences in Example D allow non-singleton indifference curves.

29. Suppose that $I = 1$, $R(\Theta) = \mathcal{P}$, $A = \{b, c, d\}$, and hence $\#(A) \geq 3$. The SCF defined by $f(\cdot) = b$ is trivially monotonic, but not dictatorial (and hence cannot be Paretian in this single-agent setting); there always exists $P \in \mathcal{P}$ such that either cPb or dPb. It can be implemented by any game form (g, S) in dominant strategies with $g(\cdot) = b$.

30. This example shows that monotonicity is not a necessary condition for an SCC to be *implementable* in dominant strategies, but it is indeed a

necessary condition for an SCC to be *fully implementable* in dominant strategies.

Suppose that $A = \{a, b, c\}$ (so that $\#(A) \geq 3$), $I = 1$, $R(\Theta) = \{R, R'\}$, $bP'aP'c$ and $bPcPa$. The SCC defined by $f(R) = \{a, b\}$ and $f(R') = \{b\}$ is implemented trivially by any game form (g, S) with $g(\cdot) = b$. This SCC is not Paretian, as a is allowed in state R, even though bPa. This SCC is not monotonic either: when moving from R to R', a moves up the agent's ranking, but $a \in f(R) \setminus f(R')$! However, it is easy to see that for full implementability, monotonicity is indeed required.

31. In fact, the proviso f being Paretian can be removed from Theorem D3, which becomes Theorem D4.

Theorem D4. (Gibbard 1973; Sattherthwaite 1975) Suppose that $\mathcal{P}^I \subset R(\Theta)$ and $\#(A) \geq 3$. Then, any SCF that is implementable in dominant strategies is dictatorial.[2]

As being dictatorial, f must be Paretian, an immediate consequence of Theorems D4 is that, with $\mathcal{P}^I \subset R(\Theta)$ and $\#(A) \geq 3$, any SCF that is implementable in dominant strategies is Paretian.

32. We now turn the spotlight to restricted preference domains. We shall first consider the space of quasi-linear utility functions, and then impose further restrictions (convex domains, convex and upper semi-continuous preferences, differentiable utility functions) on this space.

33. From now on, we shall assume $f(\Theta) = A$.[3]

The set of feasible social choices may be finite or infinite. We shall first consider the finite case. Let $\#(A) = p$, where $p \geq 3$ is an integer. The p elements of A describe the p possible states of some public good(s), from

[2] Gibbard, A., 1973, Manipulation of voting schemes: A general result, *Econometrica*, 41, 587-602.
Sattherthwaite, M.A., 1975, Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions, *Journal of Economic Theory*, 10, 187-217.

[3] It makes no sense to assume that $f(\theta) \in A^c$ for some $\theta \in \Theta$. On the other hand, given f, attention can be confined to $f(\Theta)$; we shall not be interested in elements of A that are never socially acceptable.

which, besides "money" (some private good resulting in transferable utilities among the I agents), an agent may also derive utility.

More specifically, agent i's utility is

$$u_i(x, t_i) = v_i(x) + t_i, \ \forall x \in A, \ \forall t_i \in \Re,$$

where x is the state of the public good and t_i is the amount of money held by agent i. Note that the *valuation function* $v_i(\cdot)$, a p-vector, completely describes agent i's preference. Note also that we place a subscript on t but not on x, indicating the distinction between public goods and private goods.

34. Let $V = \Re^{Ip}$ be the space of all profiles $(v_1(\cdot), v_2(\cdot), \cdots, v_I(\cdot))$. An SCC f is a correspondence $f : V \to A$.[4] An SCF is also referred to as a decisive SCC; a *generically decisive* SCC f is an SCC with the property that there exists an open dense subset V^* of V (where the usual topology is assumed) such that $f(v)$ is a singleton for all $v \in V^*$.

35. A mechanism in the current set-up is a tuple (g, M, t), where $M = \Pi_{i=1}^{I} M^i$ is the space of feasible message profiles, $g : M \to A$ specifies an element in A for each feasible message profile, and $t : M \times I \to \Re$ specifies for each agent i a monetary transfer which is again contingent on the message profile $m \in M$.

36. The mechanism (g, M, t) is a Clarke-Groves mechanism if (i) it is direct (i.e. $M = V$); (ii) $\forall v \in V$, $g(v) \in \text{argmax}_{y \in A} \sum_{i \in I} v_i(y) - c(y)$ (i.e. the decision always attains productive efficiency, where $c(y)$ is a cost function of y);[5] and (iii) $\forall i \in I$, $\forall v \in V$, $t(v, i) = h_i(v_{-i}) + \sum_{j \neq i} v_j(g(v))$ for some mapping $h_i : V_{-i} \to \Re$.[6]

    It is easy to see that if (g, M, t) is a Clarke-Groves mechanism,[7] then it truthfully implements the SCC g.

[4] Note that here the central planner does not express preferences about t (according to the SCC he is endowed with). Things will be different if he does.

[5] Most authors considered the case where $c(y) \equiv 0$.

[6] In (ii), $g(v)$ maximizes a *social welfare function* that is *utilitarian*; see Chapter 22 of Mas-Colell, Whinston, and Green (1995) for the definition of social welfare functions and their properties. A utilitarian social welfare function assumes that it is possible to make interpersonal utility comparisons.

[7] Clarke, E.H., 1971, Multipart pricing of public goods, *Public Choice*, 8, 19-33. Groves, T., 1973, Incentives in teams, *Econometrica*, 41, 617-631.

37. Consider an economy with I agents and a central planner, facing the provision of some public good (a bridge, to be concrete). Building the bridge costs c. The set A of feasible social choices is $\{0, 1\}$, where 1 means to build the bridge, and 0 means not to. Preferences are (not just quasi-) linear:

$$u_i(y, t_i) = \theta_i y + t_i, \ \forall i \in I, \ \forall y \in A.$$

We can normalize $c = 0$ (by defining $\eta_i \equiv \theta_i + \frac{c}{I}$).

38. (**Groves SCC.**) The central planner is assumed to have been endowed with the following Groves SCC:

$$y(\theta) = 1, \quad \text{iff} \quad \sum_{i=1}^{I} \theta_i \geq 0.$$

The Groves SCC is a non-dictatorial, monotonic, Paretian SCF.

39. (**Groves Mechanism**.) Groves (1973) shows that the above SCC can be implemented in dominant strategies by the following straightforward mechanism, where $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_I)$ is a profile of agents' reports about their own θ_i's:

$$y(\hat{\theta}) = 1, \quad \text{iff} \quad \sum_{i=1}^{I} \hat{\theta}_i \geq 0,$$

$$t_i(\hat{\theta}) = \begin{cases} \sum_{j \neq i} \hat{\theta}_j + h_i(\hat{\theta}_{-i}), & \text{if } y(\hat{\theta}) = 1 \\ h_i(\hat{\theta}_{-i}), & \text{if otherwise.} \end{cases}$$

Here the only other restriction we place on t is that for all $\theta \in \Theta$,

$$\sum_{i \in I} [h_i(\theta_{-i}) + y(\theta) \sum_{j \neq i} \theta_j] \leq 0,$$

so that implementing the SCC is budget feasible. Note that the Groves mechanism assumes that agents' IR conditions can be ignored, possibly because participation in this mechanism can be made mandatory.

40. To see that the Groves mechanism truthfully implements the Groves SCF, note that lying (reporting η_i in state θ_i, say) can make a difference for agent i given any $\hat{\theta}_{-i}$ only if

$$y(\eta_i, \hat{\theta}_{-i}) \neq y(\theta_i, \hat{\theta}_{-i}).$$

We take cases.

Case 1: $\theta_i + \sum_{j \neq i} \hat{\theta}_j \geq 0$.
In this case, reporting θ_i results in a payoff

$$\begin{aligned} &\theta_i \cdot y(\theta_i, \hat{\theta}_{-i}) + t_i(y(\theta_i, \hat{\theta}_{-i})) \\ &= \theta_i \cdot 1 + \sum_{j \neq i} \hat{\theta}_j + h_i(\hat{\theta}_{-i}) \geq 0 + h_i(\hat{\theta}_{-i}) \\ &= \theta_i \cdot 0 + h_i(\hat{\theta}_{-i}) \\ &= \theta_i \cdot y(\eta_i, \hat{\theta}_{-i}) + t_i(y(\eta_i, \hat{\theta}_{-i})). \end{aligned}$$

Case 2: $\theta_i + \sum_{j \neq i} \hat{\theta}_j < 0$.
In this case, reporting θ_i results in a payoff

$$\begin{aligned} &\theta_i \cdot y(\theta_i, \hat{\theta}_{-i}) + t_i(y(\theta_i, \hat{\theta}_{-i})) \\ &= \theta_i \cdot 0 + h_i(\hat{\theta}_{-i}) = h_i(\hat{\theta}_{-i}) \geq \theta_i + \sum_{j \neq i} \hat{\theta}_j + h_i(\hat{\theta}_{-i}) \\ &= \theta_i \cdot y(\eta_i, \hat{\theta}_{-i}) + t_i(y(\eta_i, \hat{\theta}_{-i})). \end{aligned}$$

Note that if we rule out the possibility of $\theta_i = 0$ for at least one $i \in I$, preferences on A are sure to be strong. In this case, truthful implementation is equivalent to full implementation of f (as f is an SCF); cf. Theorem D4.

41. **Theorem D5.** If $\Theta = \mathcal{R}^I$, then any revelation mechanism that implements the Groves SCC is a Groves mechanism. But, in this case, the planner's budget cannot be always balanced; i.e. it is impossible that $\sum_{i=1}^{I} t^i(\theta) = 0, \ \forall \theta \in \Theta$.

Proof We first show that given the Groves SCC

$$y(\theta) = \delta_{[\underline{1}'\theta \geq 0]},$$

where δ_A is the indicator function for event A and $\underline{1}$ is an I-column vector of ones, any implementing revelation mechanism has $t^i(\theta)$ taking the form of a Groves mechanism.

Let θ and $(\theta^{i\prime}, \theta^{-i})$ be two possible states. Under the implementing mechanism, truthtelling must be a dominant strategy for every agent in every state. For this, we need:

$$(a) \quad y(\theta) = y(\theta^{i\prime}, \theta^{-i}) \Rightarrow t^i(\theta) = t^i(\theta^{i\prime}, \theta^{-i}),$$

which should be obvious. Next, suppose the SCR implies that $y(\theta) = 1$ and $y(\theta^{i\prime}, \theta^{-i}) = 0$. That is,

$$(b) \quad \theta^i \geq -\sum_{j \neq i} \theta^j \geq \theta^{i\prime}.$$

In this case, truthtelling must still be optimal for i:

$$(c) \quad \theta^i + t^i(\theta) \geq t^i(\theta^{i\prime}, \theta^{-i}), \quad \text{in state } \ \theta;$$

$$(d) \quad \theta^{i\prime} + t^i(\theta) \leq t^i(\theta^{i\prime}, \theta^{-i}), \quad \text{in state } \ (\theta^{i\prime}, \theta^{-i}).$$

From (c) and (d), we have

$$(e) \quad \theta^i \geq t^i(\theta^{i\prime}, \theta^{-i}) - t^i(\theta) \geq \theta^{i\prime}.$$

As (e) has to hold for all states θ and $(\theta^{i\prime}, \theta^{-i})$ such that (b) holds, it has to be that

$$t^i(\theta^{i\prime}, \theta^{-i}) - t^i(\theta) = -\sum_{j \neq i} \theta^j,$$

which implies that (i) the transfer to agent i in states (such as θ) where $y = 1$ differs from the transfer in states (such as $(\theta^{i\prime}, \theta^{-i})$) where $y = 0$ in the extra payment $\sum_{j \neq i} \theta^j$; and (ii) from (a), the transfer to agent i in states where $y = 0$ is independent of agent i's report $\theta^{i\prime}$: $t^i(\theta^{i\prime}, \theta^{-i}) = h^i(\theta^{-i})$ for some $h^i(\cdot)$. We conclude that the implementing revelation mechanism is a Groves mechanism.

Next, we must show that the planner's budget cannot be balanced in all states θ. It suffices to consider $I = 2$. Fix $\theta^1 > \theta^{1\prime}$ and consider θ^2 such that

$$\theta^1 + \theta^2 > 0 > \theta^{1\prime} + \theta^2,$$

or equivalently,

$$\theta^2 \in (-\theta^1, -\theta^{1\prime}) \equiv B.$$

Let us assume the existence of two functions $h^1(\cdot)$ and $h^2(\cdot)$ such that given θ^1, $\theta^{1\prime}$, and that $\theta^2 \in B$,

$$(f) \quad [\theta^2 + h^1(\theta^2)] + [\theta^1 + h^2(\theta^1)] = 0,$$

$$(g) \quad h^1(\theta^2) + h^2(\theta^{1\prime}) = 0.$$

We will see a contradiction arising from this assumption. Pick ϵ small enough so that $\theta^{2\prime}, \theta^{2\prime\prime} \in B$, where

$$\theta^{2\prime} \equiv \epsilon - \theta^1, \quad \theta^{2\prime\prime} \equiv -\epsilon - \theta^{1\prime}.$$

From budget balancedness, we have for all ϵ small enough

$$h^2(\theta^{1\prime}) - h^2(\theta^1) = \epsilon = [(-\theta^{1\prime}) - (-\theta^1)] - \epsilon,$$

which is impossible. This concludes the proof.

42. A budget-balanced implementing revelation mechanism may still exist when the SCC to be implemented is not Groves.

Consider the following SCC, which is the i-th dictatorship:

$$y^i_d(\theta) = \delta_{[\theta^i \geq 0]}.$$

The balanced transfer functions in this case can be $t^i(\theta) = 0$ for all $\theta \in \Theta$.

For nondictatorial SCCs, balancedness needs asymmetry of $y(\theta)$ in its arguments. For example, the following SCC is not a dictatorial: take $I = 3$,

$$y^*(\theta) = \delta_{[\theta^2 + \theta^3 \geq 0]}.$$

This SCC can be implemented in dominant strategy by the following balanced transfer functions: If $y^* = 1$, then

$$t^1 = -\theta^2 - \theta^3, \quad t^2 = \theta^3, \quad t^3 = \theta^2;$$

and if $y^* = 0$, then $\forall i$, $t^i = 0$.

43. From now on, further restrictions will be placed on the agents' preferences but some generalizations will be allowed on the set A.

44. Now we review Green and Laffont (1977),[8] where the set A is assumed to be a compact topological space, which includes the aforementioned Groves model as a special case.[9] On the other hand, Green and Laffont assume that V_i, for all $i \in I$, consists of all upper semi-continuous functions. It is well known that *a real-valued upper semi-continuous function defined on a compact topological space always has a maximum.*[10]

---

[8]Green, J. and J.-J. Laffont, 1977, Charaterization of satisfactory mechanisms for the revelation of preferences for public goods, *Econometrica*, 45, 427-438.

[9]There are many other possibilities: (i) (indivisibility) $A = \{0, x_1, x_2, \cdots, x_p\}$; (ii) (perfect divisibility with a threshold) $A = \{0\} \bigcup [\underline{x}, \overline{x}] \subset \Re$; (iii) (multi-project case) $A = \Pi_{l=1}^{L}[0, \overline{x}_l]$; (iv) (tax laws) A is the set of continuous functions on the unit interval $[0, 1]$.

[10]Recall that given a set A, a *topology* τ on A is a collection of subsets (called *open sets*) of A such that (i) $\emptyset, A \in \tau$; (ii) for any finite collection $\{G_1, G_2, \cdots, G_n\} \subset \tau$, $\bigcap_{j=1}^{n} G_j \in \tau$; and (iii) for any collection $\{G_j; j \in J\} \subset \tau$, where J is an arbitrary (possibly uncountably infinite) index set, $\bigcup_{j \in J} G_j \in \tau$. The pair (A, τ) is called a topological space. We call (A, τ) a compact topological space, if for any collection $\{G_j; j \in J\} \subset \tau$ with $A \subset \bigcup_{j \in J} G_j$, where again J is an arbitrary index set, there exists $\{G_j; j = 1, 2, \cdots, n\} \subset \{G_j; j \in J\}$ such that $A \subset \bigcup_{j=1,2,\cdots,n} G_j$. That is, any open covering of A has a finite subcovering. A function $s : A \to \Re$ is said to be *upper semi-continuous* if for all $r \in \Re$, the preimage $s^{-1}((-\infty, r)) \in \tau$. The function s is lower semi-continuous if $-s$ is upper semi-continuous, and s is continuous if it is both upper and lower semi-continuous. We claim that if (A, τ) is a compact topological space and $s : A \to \Re$ is upper semi-continuous, then there exists $x^* \in A$ such that $s(x^*) \geq s(a)$, $\forall a \in A$. To see this, note that $\{s^{-1}((-\infty, r)); r \in \Re\}$ is an open covering of A, and hence must have a finite subcovering. This implies that $s(A)$ is bounded above in $\Re$. Since $s(A)$ is non-empty, we know it has a supremum, denoted by, say, s^*. Define

$$F_n = \{x \in A : \ s(x) \geq s^* - \frac{1}{n}\}, \ \forall n \in \mathbf{Z}_{++}.$$

The collection $\{F_n, n \in \mathbf{Z}_{++}\}$ of closed sets possesses the so-called finite intersection property: pick any finite number of elements of this collection and the intersection of these elements will never be empty. Recall that if A is compact, then any collection of closed sets satisfying the finite intersection property must also have a non-empty intersection itself. To see this, note that since A is the universe set in question, $\{G_j : j \in J\}$ is an open covering of A if and only if $\{G_j^c : j \in J\}$ is a collection of closed sets with empty intersection. Thus A is compact if and only if any collection of closed sets with empty intersection has a finite sub-collection with empty intersection. Equivalently, A is compact if and only if for all collection $\mathcal{F} \equiv \{G_j^c : j \in J\}$ of closed sets such that any finite sub-collection of $\mathcal{F}$ has a non-empty intersection (i.e. $\mathcal{F}$ possesses the *finite intersection*

45. Now, assume with Green and Laffont (1977) that A is compact and v is a profile of u.s.c. *valuation functions*. In this new context, we define the Groves SCF: An SCF $f : \Theta \to A$ is *Groves* if and only if for all $v \in V$,

$$f(v) \in \arg\max_{y \in A} \sum_{i \in I} v_i(y).$$

46. **Theorem D6.** (Green-Laffont 1977) Suppose that A is a compact topological space and for all $i \in I$, V_i consists of all u.s.c. functions on A. Given the Groves SCF, any implementing direct mechanism is a Groves mechanism.[11]

47. The domain $V = \Pi_{i \in I} V_i$ is said to be *smoothly connected* if and only if for all $i \in I$, for all $v_i, w_i \in V_i$, for all $v_{-i} \in V_{-i}$, there exists a one-dimensional parameterized family of valuation functions in V_i, denoted

$$V_i(v_i, w_i; v_{-i}) \equiv \{v_i(\cdot; y_i) \in V_i : y_i \in [0, 1]\},$$

such that for all $a \in A$,

- $v_i(a; 0) = v_i(a)$;
- $v_i(a; 1) = w_i(a)$;
- $\frac{\partial v_i(a;y_i)}{\partial y_i}$ exists for all $y_i \in [0, 1]$;[12]
- there exists $K \in \Re_{++}$ such that for all $y_i, y_i' \in [0, 1]$,

$$\left|\frac{\partial v_i(f(v_i(\cdot, y_i'), v_{-i}), y_i)}{\partial y_i}\right| \leq K.$$

48. **Theorem D7.** (Holmstrom 1979) Given the Groves SCF, every implementing straightforward mechanism is a Groves mechanism if V is smoothly connected.[13]

---

property), the intersection of all elements of $\mathcal{F}$ is also non-empty. It follows that for some $x^* \in A$, we have

$$s^* \geq s(x^*) \geq s^* - \frac{1}{n}, \ \forall n \in \mathbf{Z}_{++} \Rightarrow s(x^*) = s^*.$$

[11]Theorem D5 remains valid if A is a compact interval in $\Re$ and V_i is restricted (from the set of all u.s.c functions on A) to contain only all continuous functions A.

[12]One-sided derivatives are meant at 0 and 1.

[13]Holmstrom, B., 1979, Groves' scheme on restricted domains, *Econometrica*, 47, 1137-1144.

49. **Theorem D8.** (Holmstrom 1979) If V is a convex domain, then V is smoothly connected and hence any straightforward mechanism truthfully implementing the Groves SCF is a Groves mechanism.

50. Examples of convex domains abound: (i) V_i consists of all u.s.c. functions on A; (ii) V_i consists of all continuous functions on A; (iii) V_i consists of all strictly concave (or concave) functions on A, where A is a convex subset of some finite-dimensional Euclidean space; (iv) V_i consists of all quadratic functions on A, where A is a convex subset of some finite-dimensional Euclidean space.

51. Without smooth connectedness, straightforward mechanisms may not be Groves. Suppose that $I = 2$, $A = [0, 1]$, and V_i's are such that they consist respectively of the following parametric functions: $\forall a \in A$,

$$\forall x_1 \in [0,1],\ v_1(a; x_1) = \begin{cases} 0, & a \leq x_1; \\ x_1 - a, & a \geq x_1; \end{cases}$$

$$\forall x_2 \in [0,1],\ v_2(a; x_2) = x_2 + \frac{a}{2}.$$

Thus the Groves SCF prescribes $f(x_1, x_2) = x_1$. We now claim that there is a non-Groves straightforward mechanism truthfully implementing the Groves SCF. To this end, let g_1, g_2 be the transfer functions specified in a Groves mechanism. Define new transfer functions

$$t_1(x) = g_1(x) + \frac{x_1}{4},\ t_2(x) = g_2(x).$$

Verify that $(f, [0,1]^2, t)$ is a non-Groves straightforward mechanism that truthfully implements f.

52. The issues of uniqueness and balancedness of implementing straightforward mechanisms for the Groves SCF are reconsidered in Laffont and Maskin (1980) in a continuously differentiable setting.[14]

More specifically, Laffont and Maskin (1980) assume that for all $i \in I$, agent i has utility function

$$u_i(x, t_i) = v_i(x, \theta_i) + t_i,$$

[14] Laffont, J.-J., and E. Maskin, 1980, A differential approach to dominant strategy mechanisms, *Econometrica*, 48, 1507-1520.

where $\theta_i \in \Theta_i$ is agent i's private information, with Θ_i being an open interval in $\Re$, and where the I functions $\{v_i(\cdot,\cdot); i \in I\}$ are common knowledge of the agents and the central planner. They also assume that (**Laffont-Maskin assumptions**) the set $A = (0, +\infty)$, and the I functions $\{v_i(\cdot,\cdot); i \in I\}$ are continuously differentiable and such that for all $\theta \in \Theta \equiv \Pi_{i\in I}\Theta_i$, there exists a continuously differentiable $f(\theta) \in A = (0, +\infty)$ with

$$\sum_{i\in I} v_i(f(\theta), \theta_i) \geq \sum_{i\in I} v_i(a, \theta_i),\ \forall a \in A.$$

53. **Theorem D9.** (Laffont and Maskin 1980) Under the above Laffont-Maskin assumptions, if (f, Θ, t) implements the Groves SCF truthfully in dominant strategy, then

$$t_i(\theta) = \sum_{j\neq i} v_j(f(\theta), \theta_j) + h_i(\theta_{-i}),$$

where $h_i(\cdot)$ is an arbitrary continuously differentiable function. That is, straightforward mechanisms must be Groves.

Proof Let $t^*(\theta)$ denote the I-transfer functions specified by the Groves mechanism. Then, if $t(\theta)$ is the I-transfer functions specified by another implementing straightforward mechanism, then they both satisfy the following strategy-proofness condition:[15]

$$\forall \theta \in \Theta,\ \frac{\partial t_i}{\partial \theta_i}(\theta) \equiv -\frac{\partial v_i}{\partial x}(f(\theta), \theta_i)\frac{\partial f}{\partial \theta_i}(\theta)$$

$$= \frac{\partial t_i^*}{\partial \theta_i}(\theta),$$

[15] A game with incomplete information is *strategy-proof* if it is a weakly-dominant strategy for every player to reveal his/her private information, i.e. one fares best or at least not worse by being truthful, regardless of what the others do. Such a game is also referred to as truthful, but it is not always immune to collusion. A stronger requirement is *group strategy-proofness*: no coalitions of players can collude to misreport their preferences in a way that makes every member of the coalition better off. See https://en.wikipedia.org/wiki/Strategyproofness.

and hence by integrating, we have

$$t_i(\theta) = t_i^*(\theta) + h_i(\theta_{-i}),$$

where $h_i(\cdot)$ is an arbitrary continuously differentiable function that does not depend on θ_i.

54. The differential approach employed by Laffont and Maskin has the merit of making it easy to find the desired transfer functions. For example, suppose that for all $i \in I$,

$$v_i(x, \theta_i) = \theta_i x - \frac{x^2}{2}.$$

In this case, the strategy-proofness condition reduces to the following differential equation:[16]

$$\frac{\partial t_i}{\partial \theta_i}(\theta) = -\frac{1}{I}[\theta_i - \frac{\sum_{j\in I}\theta_j}{I}] = -\frac{1}{I}[1 - \frac{1}{I}]\theta_i + \frac{1}{I^2}\sum_{j\neq i}\theta_j,$$

so that, after integrating, we have

$$t_i(\theta) = \frac{1}{2I}[1 - \frac{1}{I}]\theta_i^2 + \frac{1}{I^2}[\sum_{j\neq i}\theta_j]\theta_i + h_i(\theta_{-i}).$$

One can verify that $t_i(\theta)$ so obtained differs from

$$\sum_{j\neq i} v_j(f(\theta), \theta_j) = \sum_{j\neq i}[\theta_i(\frac{\sum_{k\in I}\theta_k}{I}) - \frac{1}{2I^2}(\sum_{l\in I}\theta_l)^2]$$

by simply a function of θ_{-i}.

55. **Theorem D10.** (Laffont-Maskin 1980) Under the above Laffont-Maskin assumptions, the Groves mechanism can be chosen to be balanced if and only if for all $\theta \in \Theta$,

$$\sum_{i\in I} \frac{\partial^{I-1}}{\partial\theta_1\partial\theta_2\cdots\partial\theta_{i-1}\partial\theta_{i+1}\cdots\partial\theta_I}[\frac{\partial v_i}{\partial x}(f(\theta), \theta_i)\frac{\partial f}{\partial \theta_i}(\theta)] \equiv 0.$$

[16] Note that the Groves SCF specifies $f(\theta) = \frac{\sum_{i\in I}\theta_i}{I}$, and hence $\frac{\partial f}{\partial \theta_i}(\theta) = \frac{1}{I}$.

Proof Being balanced, the transfer functions must be such that, for all $\theta \in \Theta$,

$$\sum_{i \in I} t_i(\theta) = 0,$$

or, equivalently, we write

$$\sum_{i \in I} t_i(\theta) \equiv 0.$$

Using the strategy-proofness condition, we thus have

$$\sum_{i \in I} [-\int \frac{\partial v_i}{\partial x} \frac{\partial f}{\partial \theta_i} + h_i(\theta_{-i})] \equiv 0.$$

Differentiating this identity with respect $\theta_1, \theta_2, \cdots, \theta_I$, we obtain the condition. This establishes necessity.

For sufficiency, we integrate

$$\sum_{i \in I} \frac{\partial^{I-1}}{\partial \theta_1 \partial \theta_2 \cdots \partial \theta_{i-1} \partial \theta_{i+1} \cdots \partial \theta_I} [\frac{\partial v_i}{\partial x}(f(\theta), \theta_i) \frac{\partial f}{\partial \theta_i}(\theta)] \equiv 0$$

with respect to $\theta_1, \theta_2, \cdots, \theta_I$ and work backwards. We obtain

$$\sum_{i \in I} [-\int \frac{\partial v_i}{\partial x} \frac{\partial f}{\partial \theta_i} + h_i(\theta_{-i})] \equiv 0.$$

Using the strategy-proofness condition, we thus have

$$\sum_{i \in I} t_i(\theta) \equiv 0.$$

That is, the mechanism is balanced.

56. The merit of Theorem D10 is that it allows us to infer from the primitives whether or not balanced mechanisms are possible. For example, if $I \geq 3$, then the case where v_i's are from quadratic class admits a balanced Groves mechanism: for all $i \in I$,

$$\frac{\partial v_i}{\partial x}(f(\theta), \theta_i) \frac{\partial f}{\partial \theta_i}(\theta) = -\frac{1}{I}[1 - \frac{1}{I}]\theta_i + \frac{1}{I^2} \sum_{j \neq i} \theta_j,$$

and hence given $I \geq 3$, for all i,

$$\sum_{i \in I} \frac{\partial^{I-1}}{\partial \theta_1 \partial \theta_2 \cdots \partial \theta_{i-1} \partial \theta_{i+1} \cdots \partial \theta_I} [\frac{\partial v_i}{\partial x}(f(\theta), \theta_i) \frac{\partial f}{\partial \theta_i}(\theta)] \equiv 0.$$

57. Part 2. **Nash Implementation**.

58. Suppose that $\Theta_i = \Theta_j \equiv \Theta$, for all $i, j \in \{1, 2, \cdots, I\}$. Given a game form (g, S), let $s^*(\theta)$ be one pur-strategy Nash equilibrium of the game (g, S, θ). Let $E_g(\theta)$ be the set of all pure-strategy Nash equilibria in state θ. Then, define $g(E_g(\theta)) \equiv \{g(s^*) : s^* \in E_g(\theta)\}$. We say that (g, S) fully implements an SCC f in Nash equilibrium if and only if $g(E_g(\theta)) = f(\theta)$ for all $\theta \in \Theta$, and in the presence of such a game form (g, S), we say that f is Nash implementable.

59. **Theorem N1.** (Maskin 1977; Maskin 1999) f is Nash implementable only if f is monotonic.[17]

Proof. Define the lower contour set at a for agent i in state θ by

$$L^i(a, \theta) \equiv \{b \in A : aR^i(\theta)b\}.$$

We shall prove Theorem N1 by contraposition. Recall from section 17 that f is not monotonic if and only if there exist $\theta, \phi \in \Theta$ and $a \in A$ such that for all $i = 1, 2, \cdots, I$,

$$L^i(a, \theta) \subset L^i(a, \phi),$$

and yet $a \in f(\theta) \setminus f(\phi)$.[18] We show that in this case no game forms (g, S) can fully implement f in Nash strategy.

Suppose instead that there were such a game form (g, S). Then there exists some Nash equilibrium s^* for the game (g, S, θ) such that $g(s^*) = a \in f(\theta)$. That is, for agent i, given his rival agents would play s^*_{-i} in state θ,

$$a = g(s^*_i, s^*_{-i})R^i(\theta)g(s_i, s^*_{-i}),\ \forall s_i \in S^i,$$

[17] Maskin, E., 1977, Nash Equilibrium and Welfare Optimality, MIT working paper. Maskin, E., 1999, Nash Equilibrium and Welfare Optimality, *Review of Economic Studies*, 66, 23-38.

[18] An equivalent definition for f being monotonic is this: for any $\theta, \phi \in \Theta$ and $a \in A$ with $a \in f(\theta) \setminus f(\phi)$, there must exist agent i and some $b \in A$ such that $b \in L^i(a, \theta) \setminus L^i(a, \phi)$, or such that $aR^i(\theta)b$ but $bP^i(\phi)a$.

$$\Rightarrow g(s_i, s^*_{-i}) \in L^i(a,\theta) \subset L^i(a,\phi), \ \forall s_i \in S^i,$$

but then s^*_i continues to be agent i's best response against s^*_{-i} in state ϕ! As this is true for all agents i, we conclude that s^* would also arise as a pure-strategy Nash equilibrium in state ϕ. But then $a \in g(E_g(\phi)) \setminus f(\phi)$, showing that (g, S) does not fully implement f in Nash equilibrium.

60. To state our next result, we introduce the notion of no veto power. An SCC f satisfies (weak) no veto power if for all $i \in \{1, 2, \cdots, I\}$, for all $\theta \in \Theta$, and for all $a \in A$,

$$L^j(a,\theta) = A, \forall j \neq i \Rightarrow a \in f(\theta).$$

In words, if a is top ranked by all agents $j \neq i$ in state θ, then $a \in f(\theta)$ whether agent i likes a or not. (Agent i has no veto power!)

61. **Theorem N2.** (Maskin 1977; Repullo 1987[19]) Suppose that $I \geq 3$, and that f is monotonic and satisfies no veto power. Then f is Nash implementable.

Proof. The proof is by construction of a canonical game form (g, S) which fully implements f. Define for all i, $S^i = \Theta \times A \times \mathbf{Z}_+$, where $\mathbf{Z}_+$ denotes the set of positive integers, and define $g : S \to A$ as follows:

(a) If s is such that there exists $i \in \{1, 2, \cdots, I\}$ such that $s_i = (\eta, a_i, k_i)$ and for all $j \neq i$, $s_j = (\theta, a, k)$ with $a \in f(\theta)$, then

$$g(s) = \begin{cases} a_i, & \text{if } a_i \in L^i(a,\theta); \\ a, & \text{otherwise.} \end{cases}$$

(b) If s is such that (a) does not apply, then $g(s) = a_i$ where i is an agent announcing the highest k_i, with ties being broken by selecting among the agents announcing the highest k_i the person with the smallest i.

[19] Repullo, R., 1987, A Simple Proof of Maskin's Theorem on Nash Implementation, *Social Choice and Welfare*, 4, 39-41.

We shall show first that $f(\theta) \subset g(E_g(\theta))$ for all $\theta \in \Theta$, and then that $g(E_g(\theta)) \subset f(\theta)$ for all $\theta \in \Theta$.

- $f(\theta) \subset g(E_g(\theta))$ for all $\theta \in \Theta$.

  Given any $a \in f(\theta)$, define for all i, $s_i = (\theta, a, 1)$. Then s is such that (a) holds, and if agent i alone would like to deviate and to implement another a_i, he must choose some $a_i \in L^i(a, \theta)$, and hence he has no incentive to make unilateral deviations. Thus $a \in g(E_g(\theta))$, and this being true for all $\theta \in \Theta$ and for all $a \in f(\theta)$, we conclude that $f(\theta) \subset g(E_g(\theta))$ for all $\theta \in \Theta$.
- $g(E_g(\theta)) \subset f(\theta)$ for all $\theta \in \Theta$

  Let $s \in E_g(\theta)$, and we shall show that $g(s) \in f(\theta)$. Suppose θ is the true state. We take cases.
  - Suppose that s is such that $s_i = (\eta, a, k)$ $\forall i \in \{1, 2, \cdots, I\}$, with $a \in f(\eta)$, so that $g(s) = a$.

    For all $i \in \{1, 2, \cdots, I\}$, if agent i wishes to deviate unilaterally from s, then according to (a) above, agent i must announce some $s'_i = (\phi, a_i, k_i)$ with $a_i \in L^i(a, \eta)$. Since s is a Nash equilibrium in the true state θ, agent i weakly prefers the equilibrium outcome $a = g(s)$ to a_i in the true state θ, and this implies that

    $$a_i \in L^i(a, \eta) \Rightarrow a_i \in L^i(a, \theta),$$

    and this being true for all $i \in \{1, 2, \cdots, I\}$, we conclude that $a \in f(\theta)$ since f is monotonic.
  - Suppose that s is such that $s_i = (\eta, a, k)$ $\forall i \in \{1, 2, \cdots, I\}$, with $a \notin f(\eta)$.

    In this case, by (b), any agent i can deviate and announce $s'_i = (\phi, a_i, k')$, where $k' > k$, so that the outcome a_i rather than

$g(s)$ would be implemented. Since s is a Nash equilibrium in the true state θ, it must be that

$$g(s)R^i(\theta)a_i, \ \forall a_i \in A,$$

or equivalently,

$$L^i(g(s), \theta) = A,$$

and with this being true for each single agent i, we conclude that $g(s) \in f(\theta)$ by the fact that f satisfies no veto power.

– Suppose that s is such that there exist $i \neq j$, $s_i \neq s_j$.

  In this case, thanks to the fact that $I \geq 3$, some agent $h \notin \{i, j\}$ can implement any $a_h \in A$ by announcing an integer k_h exceeding k_n for all $n \neq h$. Since s is a Nash equilibrium in the true state θ, it must be that

$$L^h(g(s), \theta) = A, \ \forall h \notin \{i, j\}.$$

  Moreover, it is impossible that $s_h = s_i$ and $s_h = s_j$, simply because $s_i \neq s_j$. Suppose that $s_h \neq s_i$. Then we can repeat the above argument and conclude that

$$L^j(g(s), \theta) = A.$$

  It follows that $g(s)$ is top ranked in state θ by all agents $n \neq i$, so that $g(s) \in f(\theta)$ by the fact that f satisfies no veto power.

62. Moore and Repullo (1990) provide equivalence conditions for Nash implementable SCCs.[20] We shall consider the case where $I \geq 3$ in section 63, and the case where $I = 2$ in section 64.

63. Suppose that $I \geq 3$ and (g, S) fully implements an SCC f. We shall derive three necessary conditions.

    Let $M_i(D, \theta)$ denote the set of agent i's top-ranked elements in $D \subset A$ in state θ.

[20] Moore, J., and R. Repullo, 1990, Nash Implementation: A Full Characterization, *Econometrica*, 58, 1083-1099.

Define

$$B = \{a \in A : a = g(s) \text{ for some } \ s \in S\}.$$

Hence B is the image set of $g(\cdot)$.

Given $\theta \in \Theta$ and $a \in f(\theta)$, let $s(a,\theta)$ be one Nash equilibrium for g in state θ, and given $s(a,\theta)$, define accordingly

$$C_i(a,\theta) = \{c \in A : c = g(s'_i, s_{-i}(a,\theta)) \text{ for some } \ s'_i \in S_i\}.$$

Thus $C_i(a,\theta)$ is the set of outcomes that agent i can generate by varying his own move given that the other agents are playing $s_{-i}(a,\theta)$.

Note first that

$$a \in \bigcap_{i=1}^{I} M_i(C_i(a,\theta),\theta^*) \Rightarrow a \in f(\theta^*). \tag{1}$$

This happens because when $a \in \bigcap_{i=1}^{I} M_i(C_i(a,\theta),\theta^*)$, $s(a,\theta)$ must also be an NE in state θ^*, and since we have assumed that g fully implements f, $s(a,\theta)$ cannot be an unwanted equilibrium in state θ^*.

Next, suppose that $c = g(s'_i, s_{-i}(a,\theta))$ for some $s'_i \in S_i$. Then the following assertion must be true:

$$c \in M_i(C_i(a,\theta),\theta^*) \bigcap [\bigcap_{j \neq i} M_j(B,\theta^*)] \text{ for some } \ i \Rightarrow c \in f(\theta^*). \tag{2}$$

Indeed, if $c \in M_i(C_i(a,\theta),\theta^*)$ then in state θ^*, s'_i is one optimal choice for agent i if agent i alone would like to deviate from $s(a,\theta)$. Moreover, expecting agent i's deviation move s'_i, the remaining agents would still stick to $s_{-i}(a,\theta)$ if c is one top-ranked choice in B from each and every agent $j \neq i$. In this case, the strategy profile $(s'_i, s_{-i}(a,\theta))$ is one equilibrium in state θ^* under g. Since (g,S) fully implements f, $(s'_i, s_{-i}(a,\theta))$ cannot be an unwanted equilibrium in state θ^*. Thus we conclude that

$$c = g(s'_i, s_{-i}(a,\theta)) \in f(\theta^*).$$

Finally, if d in a top-ranked element in B in state θ^* from each and every agent's perspective, then obviously any strategy profile contained in the pre-image $g^{-1}(d)$ is one Nash equilibrium in state θ^*. That is, we must have

$$d \in \bigcap_{i=1}^{I} M_i(B, \theta^*) \Rightarrow d \in f(\theta^*). \quad (3)$$

Condition μ. There exists a set $B \subset A$, and for each $i \in \{1, 2, \cdots, I\}$, each $\theta \in \Theta$, and each $a \in f(\theta)$, there exists a set $C_i(a, \theta) \subset B$ with $a \in M_i(C_i(a, \theta), \theta)$ such that for all $\theta^* \in \Theta$, the above (1), (2), and (3) are satisfied.

Theorem N3. (Moore and Repullo 1990) If $I \geq 3$, then f is Nash implementable if and only if Condition μ is satisfied.

Proof. The necessity has been outlined above. If (g, S) fully implements f, then by defining B and $C_i(a, \theta)$ as above, we see that $a \in M_i(C_i(a, \theta), \theta)$, and moreover, for all $\theta^* \in \Theta$, the above (1), (2), and (3) are satisfied.

Now to prove sufficiency, we assume that B and the sets $C_i(a, \theta)$ exist and satisfy Condition μ, and we shall construct a game form (g, S) that fully implements f.

Let N be the set of non-negative integers. Define for all $i \in \{1, 2, \cdots, I\}$,

$$S_i \equiv \{(\theta_i, a_i, b_i, n_i) \in \Theta \times A \times B \times N : \ a_i \in f(\theta_i)\}.$$

Define g accordingly as follows.

- (A). If s is such that $s_i = (\theta, a, b, n)$ for all $i \in \{1, 2, \cdots, I\}$, then $g(s) = a$.
- (B). If there exists i such that $s_j = (\theta, a, b, n) \neq s_i$ for all $j \neq i$, then

$$g(s) = \begin{cases} b_i, & \text{if } \ b_i \in C_i(a, \theta); \\ a, & \text{otherwise.} \end{cases}$$

- (C). If neither (A) nor (B) applies, then we let $g(s) = b_i$, where i is the lowest-indexed agent among those annoucing the highest integer n_i.

We must show $f(\theta) \subset g(E_g(\theta))$ and $g(E_g(\theta)) \subset f(\theta)$, for all $\theta \in \Theta$.

To show that $f(\theta) \subset g(E_g(\theta))$ for all $\theta \in \Theta$, pick any $a \in f(\theta)$ and let $s_i = (\theta, a, a, 0)$ for all $i \in \{1, 2, \cdots, I\}$. By (A) above, $g(s) = a$ if nobody deviates from s. An individual agent i alone can implement some outcome $b_i \in C_i(a, \theta)$, but since Condition μ holds, which implies that $a \in M_i(C_i(a, \theta), \theta)$, and hence agent i has no reason to replace a by b_i. Thus no unilateral deviations would occur, and we conclude that $a \in g(E_g(\theta))$ as required.

Next, we show that (g, S) generates no unwanted equilibria; that is, we show that $g(E_g(\theta)) \subset f(\theta)$, for all $\theta \in \Theta$.

To this end, let θ^* be the true state, and let s be one Nash equilibrium under g in state θ^*. We must show that $g(s) \in f(\theta^*)$. There are three possibilities to consider.

- Suppose that s is such that (A) holds. Then $g(s) = a$. In this case, each agent i alone can deviate unilaterally to implement an outcome contained in $M_i(C_i(a, \theta), \theta^*)$, but since s is a Nash equilibrium in state θ^*, we know that no agents would actually do that. This implies that

$$g(s) = a \in \bigcap_{i=1}^{I} M_i(C_i(a, \theta), \theta^*),$$

  and since Condition μ holds (implying that (1) holds), we must have

$$g(s) = a \in f(\theta^*).$$

- Suppose that s is such that (B) holds, and suppose that all agents other than i adopt the same strategy (θ, a, b, n). In this case, given the other $I - 1$ agents' strategies, agent i alone can implement his state-θ^* top-ranked outcomes in $C_i(a, \theta)$. Since s is an equilibrium in state θ^*, $g(s)$ must be a state-θ^* top-ranked outcome in $C_i(a, \theta)$ for agent i; i.e.,

$$g(s) \in M_i(C_i(a, \theta), \theta^*).$$

Moreover, since $I \geq 3$, each agent $j \neq i$ alone can deviate unilaterally to implement his top-ranked outcomes in B, according to (C). Since s is an equilibrium, $g(s)$ must already be one top-ranked outcome in B from the perspective of all agents other i; i.e.,

$$g(s) \in \bigcap_{j \neq i} M_j(B, \theta^*).$$

Since Condition μ holds (implying that (2) holds), we conclude that

$$g(s) \in M_i(C_i(a, \theta), \theta^*) \bigcap \bigcap_{j \neq i} M_j(B, \theta^*) \Rightarrow g(s) \in f(\theta^*).$$

- Finally, suppose that s satisfies (C). In this case each agent i alone can implement his state-θ^* top-ranked outcomes by announcing some integer $n_i > n_j$ for all $j \neq i$. Since s is an equilibrium, no agents would do that, and this implies that

$$g(s) \in \bigcap_{j=1}^{I} M_j(B, \theta^*) \Rightarrow g(s) \in f(\theta^*),$$

since Condition μ holds (implying that (3) holds).

The proof is complete.

64. In contracting problems, $I = 2$ is the leading case. So Moore and Repullo (1990) also give an equivalence condition for Nash implementation with two agents.

Suppose that $I = 2$ and (g, S) fully implements f. We shall derive a necessary condition. For $\theta, \phi \in \Theta$, $a \in f(\theta)$, and $b \in f(\phi)$, let $(s_1(b,\phi), s_2(b,\phi))$ be one equilibrium under g in state ϕ, and let $(s_1(a,\theta), s_2(a,\theta))$ be one equilibrium under g in state θ, such that

$$g(s_1(a,\theta), s_2(a,\theta)) = a, \quad g(s_1(b,\phi), s_2(b,\phi)) = b.$$

Let $e \equiv g(s_1(b,\phi), s_2(a,\theta))$; that is, e is the outcome resulting from agent 1 playing $s_1(b,\phi)$ and agent 2 playing $s_2(a,\theta)$.

Let $C_1(a,\theta)$ be the set of outcomes agent 1 can generate by varying his own move, assuming that agent 2 is playing $s_2(a,\theta)$; and $C_2(b,\phi)$ the set of outcomes agent 2 can generate by varying his own move, assuming that agent 1 is playing $s_1(b,\phi)$.

Consider a true state θ^* where

$$e \in M_1(C_1(a,\theta), \theta^*) \bigcap M_2(C_2(b,\phi), \theta^*).$$

In such a state θ^*, the strategy profile $(s_1(b,\phi), s_2(a,\theta))$ is obviously a Nash equilibrium under g, and since by assumption (g, S) fully implements f, this equilibrium cannot be an unwanted equilibrium in state θ^*. Thus we conclude that the following is a necessary condition for (g, S) to fully implements f:

$$e \in M_1(C_1(a,\theta), \theta^*) \bigcap M_2(C_2(b,\phi), \theta^*) \Rightarrow e \in f(\theta^*). \tag{4}$$

Condition $\mu 2$. Condition μ holds, and moreover, for each 4-tuple $(a,\theta,b,\phi) \in A \times \Theta \times A \times \Theta$ with $a \in f(\theta)$ and $b \in f(\phi)$, there exists some element $e = e(a,\theta,b,\phi)$ contained in $C_1(a,\theta) \bigcap C_2(b,\phi)$ such that for all $\theta^* \in \Theta$, (4) holds.

Theorem N4. (Moore and Repullo 1990) If $I = 2$, then f is Nash implementable if and only if Condition $\mu 2$ is satisfied.

Proof. Suppose that (g, S) fully implements f. Then let B and the sets $C_j(a,\theta)$ be as defined prior to expressions (1)-(3). Then we have shown that (1)-(4) must hold. This establishes necessity.

To prove sufficiency, we assume that Condition $\mu 2$ holds and construct a game form (g, S) to fully implement f.

Let N be the set of non-negative integers. Define for all $i \in \{1, 2\}$,

$$S_i \equiv \{(\theta_i, a_i, b_i, n_i) \in \Theta \times A \times B \times N : \ a_i \in f(\theta_i)\}.$$

Define g accordingly as follows.

- (A). If s is such that $(\theta_1, a_1) = (\theta_2, a_2) = (\theta, a)$, then $g(s) = a$.
- (B). If $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_1 = n_2 = 0$, then $g(s) = e(a_2, \theta_2, a_1, \theta_1)$.
- (C). If $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_1 > n_2 = 0$, then

$$g(s) = \begin{cases} b_1, & \text{if } \ b_1 \in C_1(a_2, \theta_2); \\ \\ e(a_2, \theta_2, a_1, \theta_1), & \text{otherwise.} \end{cases}$$

- (D). If $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_2 > n_1 = 0$, then

$$g(s) = \begin{cases} b_2, & \text{if } \ b_2 \in C_2(a_1, \theta_1); \\ \\ e(a_2, \theta_2, a_1, \theta_1), & \text{otherwise.} \end{cases}$$

- (E). If $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_1 \geq n_2 > 0$, then $g(s) = b_1$.
- (F). If $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_2 > n_1 > 0$, then $g(s) = b_2$.

Now we show that $f(\theta) \subset g(E_g(\theta))$ and $g(E_g(\theta)) \subset f(\theta)$, for all $\theta \in \Theta$. To show that $f(\theta) \subset g(E_g(\theta))$, for all $\theta \in \Theta$, pick any $\theta \in \Theta$ and any $a \in f(\theta)$. Define for $i = 1, 2$,

$$s_i = (\theta, a, b_i, 0),$$

so that by (A) $g(s) = a$ would be implemented unless some agent i makes a unilateral deviation, where assume without loss of generality that $i = 1$. According to g, agent 1 alone can deviate and implement a different element in A only if agent 1 announces some $s' = (\theta', a', b', n')$ with $(\theta', a') \neq (\theta, a)$, and since $n_2 = 0$, agent 1's deviation would result in (s', s_2) satisfying either (B) or (C). Note that $e(a_2, \theta_2, a_1, \theta_1) \in C_1(a_2, \theta_2)$ so that the optimal s' must satisfy (C) (i.e., s' must have $n' > 0$), implying that agent 1 would be able to implement any element in $C_1(a, \theta)$. Since θ is the true state, and since by Condition $\mu 2$ we have $a \in M_1(C_1(a, \theta), \theta)$, no unilateral deviations would make agent 1 better off. Applying the same reasoning to agent 2, we conclude that $a \in g(E_g(\theta))$. As we have arbitrarily picked a and θ, we conclude that $f(\theta) \subset g(E_g(\theta))$, for all $\theta \in \Theta$.

Now, to show that $g(E_g(\theta)) \subset f(\theta)$, for all $\theta \in \Theta$, we assume that $s \in E_g(\theta^*)$, and we must show that $g(s) \in f(\theta^*)$, where θ^* denotes the true state. There are four possibilities to consider: either s satisfies (A), or s satisfies (B), or s satisfies (C) or (D), or s satisfies (E) or (F).

- Suppose that $s = (s_1, s_2)$ satisfies (A); that is, $\theta_i = \theta,\ a_i = a$, for $i \in \{1, 2\}$.

  - First suppose that $n_1 = n_2 = 0$.

    First focus on agent 1's incentives to deviate alone from s. If agent 1 would deviate alone, then he can result in (s_1', s_2) satisfying either (B) or (C), and it would be optimal for agent 1 to implement some element in $M_1(C_1(a_2, \theta_2), \theta^*)$. Since agent 1 wishes to stick to s, it must be that $g(s) \in M_1(C_1(a_2, \theta_2), \theta^*)$. By the same reasoning, we must have $g(s) \in M_2(C_2(a_1, \theta_1), \theta^*)$. Since Condition $\mu 2$ holds, we have by (1),

    $$g(s) \in \bigcap_{i=1}^{2} M_i(C_i(a, \theta), \theta^*) \Rightarrow g(s) \in f(\theta^*).$$

  - Next, suppose that $n_i > 0$ for exactly one agent i.

    Then agent i alone can deviate and announce some $(\theta', a') \neq (\theta, a)$, and it would be optimal for agent i to maintain $n_i > 0$, because, once again, $e(a, \theta, a', \theta') \in C_i(a, \theta)$. Since agent i does not really want to deviate alone, we must have $g(s) \in M_i(C_i(a, \theta), \theta^*)$.

Now, refer to the other agent as agent j. Agent j can deviate alone to impose either (E) or (F), so that he can implement any element in B. Since agent j does not really want to deviate, we must have $g(s) \in M_j(B, \theta^*)$.
Since Condition $\mu 2$ holds, we have by (2),

$$g(s) \in M_i(C_i(a, \theta), \theta^*) \bigcap M_j(B, \theta^*) \Rightarrow g(s) \in f(\theta^*).$$

– Finally, suppose that $n_1, n_2 > 0$.

In this case, each agent i alone can deviate to impose either (E) or (F), so that he can implement any element in B. Since s is an equilibrium, no agents alone would really want to deviate, and we must conclude that, by Condition $\mu 2$, and (3),

$$g(s) \in \bigcap_{i=1}^{2} M_i(B, \theta^*) \Rightarrow g(s) \in f(\theta^*).$$

- Now, suppose instead that s satisfies (B); that is, $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_1 = n_2 = 0$, and $g(s) = e(a_2, \theta_2, a_1, \theta_1)$.

Consider agent 1's incentives to deviate alone. By announcing $(\theta_1', a_1') = (\theta_2, a_2)$, agent 1 can implement a_2. However, agent 1 can make a better deviation: since $a_2 \in C_1(a_2, \theta_2)$, agent 1 can impose (C) and implement any element of $C_1(a_2, \theta_2)$ simply by announcing $n_1' > 0$. Since agent 1 does not really want to make such a deviation, we must have

$$g(s) = e(a_2, \theta_2, a_1, \theta_1) \in M_1(C_1(a_2, \theta_2), \theta^*).$$

Upon applying the same reasoning to agent 2, we conclude by Condition $\mu 2$, and (4), that

$$e(a_2, \theta_2, a_1, \theta_1) = g(s) \in M_1(C_1(a_2, \theta_2), \theta^*) \bigcap M_2(C_2(a_1, \theta_1), \theta^*) \Rightarrow g(s) \in f(\theta^*).$$

- Now, suppose instead that s satisfies (C).

That is, s is such that $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_1 > n_2 = 0$, with

$$g(s) \in M_1(C_1(a_2, \theta_2), \theta^*).$$

If agent 1 wants to deviate alone, he can announce $(\theta_1', a_1') = (\theta_2, a_2)$ to implement a_2, but since $a_2 \in C_1(a_2, \theta_2)$, this deviation would not benefit agent 1. On the other hand, if agent 2 alone wants to deviate, he can implement a_1 by announcing $(\theta_2', a_2') = (\theta_1, a_1)$, or announcing some $n_2' > n_1$ to select any element in B. Recall that $a_1 \in C_2(a_1, \theta_1) \subset B$, agent 2 would focus on the latter type of deviations. Since agent 2 does not really want to deviate from s, we conclude that

$$g(s) \in M_2(B, \theta^*).$$

Thus it follows from Condition $\mu 2$, or from (2) alone, that $g(s) \in f(\theta^*)$.

Note that the case that s satisfies (D) can be analogously proved.

- Suppose that s satisfies (E).

That is, s is such that $(\theta_1, a_1) \neq (\theta_2, a_2)$ and $n_1 \geq n_2 > 0$ with

$$g(s) \in M_1(B, \theta^*).$$

Apparently, no unilateral deviations can benefit agent 1. On the other hand, if agent 2 alone wants to deviate, then he can announce some $n_2' > n_1$ and then select any element in B. Since agent 2 does not really want to deviate from s, we must have

$$g(s) \in M_2(B, \theta^*).$$

It thus follows from Condition $\mu 2$, or (3) specifically, that

$$g(s) \in \bigcap_{i=1}^{2} M_i(B, \theta^*) \Rightarrow g(s) \in f(\theta^*).$$

Note that the case where s satisfies (F) can then be analogously proved.

65. Part 3. **Subgame-perfect Implementation.**

Moore and Repullo (1988) show that an SCC that cannot be fully implemented by a static game form in Nash equilibrium may still be fully implemented by some multi-stage game form g in subgame-perfect Nash equilibrium.[21]

We first state their necessary condition:

Condition C: For each pair of preference states $\theta, \phi \in \Theta$, and for each $a \in f(\theta) \setminus f(\phi)$, there exists a finite sequence $\sigma(\theta, \phi; a) = \{a = a_0, a_1, a_2, \cdots, a_l, a_{l+1}\}$ in A with $l \geq 1$ and l is uniformly bounded above by some positive integer $\bar{l}$, such that
(i) for each $k = 0, 1, \cdots, l$, there exists a particular agent $j(k)$ satisfying $a_k R^{j(k)}(\theta) a_{k+1}$; and
(ii) there exists a particular agent $j(l)$ satisfying $a_{l+1} P^{j(l)}(\phi) a_l$.

66. Note that Condition C reduces to monotonicity of f (as defined in section 17) when $l = 1$. Here, preference reversal across states θ and ϕ does not necessarily occur between a and another outcome $b \in A$. It occurs between a_l and a_{l+1}, although we do need a to associate with a_l in a particular way (as stated in (i) of Condition C).

67. **Theorem SP1.** If f is fully implemented by some finite stage game form g, then Condition C is satisfied.

Proof. We suppose for simplicity that $I = 2$, and the argument generalizes obviously for $I > 2$.

Consider two preference states θ and ϕ and an outcome $a \in f(\theta) \setminus f(\phi)$. Since g implements fully f, there must be some SPNE s_a in state θ such that $g(s_a) = a$. Now, consider the first stage of g, where, as their first-stage actions, assume that agent 1 would choose among a finite number of rows, and agent 2 would choose among a finite number of columns. The i-th row and the j-th column then define a distinct second-stage

[21] Moore, J., and R. Repullo, 1988, Subgame Perfect Implementation, *Econometrica*, 56, 1191-1220.

subgame g_{ij}. Then $g_{ij}(s_a)$ denotes the SPNE outcome for the subgame g_{ij} when the two agents continue to adopt the strategy profile s_a.

Since a is an SPNE outcome under g in state θ, a must appear as $g_{ij}(s_a)$ for some i, j. Assume without loss of generality that $i = 1 = j$. Then we ask: when preference state has changed into ϕ from θ, would $g_{1j}(s_a)$ continue to be an SPNE outcome for the second-stage subgame g_{1j} in state ϕ? Would $g_{i1}(s_a)$ continue to be an SPNE outcome for the second-stage subgame g_{i1} in state ϕ? If the answer is positive for all i and all j, then we are done: since $g_{11}(s_a) = a$ is an SPNE outcome in state θ but not in state ϕ (simply because, by assumption, g fully implements f and $a \in f(\theta) \setminus f(\phi)$), there must exist some $(i, 1)$ with $g_{11}(s_a)R^1(\theta)g_{i1}(s_a)$ and $g_{i1}(s_a)P^1(\phi)g_{11}(s_a)$ or some $(1, j)$ with $g_{11}(s_a)R^1(\theta)g_{1j}(s_a)$ and $g_{1j}(s_a)P^2(\phi)g_{11}(s_a)$. In the former case, agent 1 strictly prefer the i-th row to the first row in state ϕ; and in the latter case, agent 2 strictly prefers the j-th column to the first column in state ϕ. In either case, we can define $l = 1$, $a = a_0 = a_1$, and a_2 equals either $g_{i1}(s_a)$ or $g_{1j}(s_a)$. Condition C holds.

In the opposite case, there must exist some i such that $g_{i1}(s_a)$ is an SPNE outcome of the second-stage subgame g_{i1} in state θ but not in state ϕ or some j such that $g_{1j}(s_a)$ is an SPNE outcome of the second-stage subgame g_{1j} in state θ but not in state ϕ. Suppose that the former is true. Then define $a_1 \equiv g_{i1}(s_a)$, and we have $a \equiv a_0 R^1(\theta) a_1$.

Now, we can focus on the second-stage subgame g_{i1}, assuming that in this subgame agent 1 would again choose among a finite number of rows, and agent 2 would choose among a finite number of columns. By turning the spotlight away from a (and away from the first stage of g) and onto a_1 (and onto the first stage of g_{i1}), we can repeat the above reasoning. Since by assumption g is a *finite* stage game form, this process must end somewhere. Thus we can obtain a sequence $a_0, a_1, \cdots, a_l$, with preferece reversal eventually appearing over the two outcomes (a_l, a_{l+1}) in states θ and ϕ. Again, Condition C must hold.

68. Now we focus on the case with $I \geq 3$, and state Moore and Repullo's sufficiency Theorem. To begin, let us assume that Condition C holds and define a class $\mathcal{Q}(f)$ of subsets Q of A as follows.

For each pair of states $\theta, \phi \in \Theta$, and for each $a \in f(\theta) \setminus f(\phi)$, select *one* sequence $\sigma(\theta, \phi; a)$ satisfying (i) and (ii) in Condition C. Then let Q be the union of the elements in these selected sequences. Then $\mathcal{Q}(f)$ comprises the Q's constructed from *all* possible selections. Now we can define:

Condition C$^+$: Condition C holds. Further, there exists a particular $Q^+ \in \mathcal{Q}(f)$, and some $B \subset A$ with $Q^+ \subset B$, such that the following is true for each $\theta \in \Theta$: for each agent i there exists a non-empty set $M^i(\theta)$ of state-θ top-ranked outcomes in B, such that (1) $M^i(\theta) \bigcap M^j(\theta) = \emptyset$ for all $i \neq j$; and (2) for all i, $M^i(\theta) \bigcap Q^+ = \emptyset$.

From now on, $\sigma^+(\theta, \phi; a)$ would stand for the selected sequence in Q^+.

69. **Theorem SP2.** If SCC f satisfies Condition C$^+$, and $I \geq 3$, then f can be implemented by a finite stage game form in subgame perfect Nash equilibrium.

Proof. Define a mechanism g as follows.

(a) First Stage (Stage 0): Each agent $i = 1, ..., I$ announces some triplet (θ_i, a_i, n_i^0), where $\theta_i \in \Theta$, $a_i \in f(\theta_i)$, and n_i^0 is a non-negative integer. There are three possibilities to consider:

- (0.1) If for some a and θ we have for all i, $\theta_i = \theta$, $a_i = a$, and $a \in f(\theta)$, then outcome a is implemented. STOP
- (0.2) If only $I - 1$ agents agree on θ and a with $a \in f(\theta)$, and if the remaining agent i announces a profile ϕ, and
  - (0.2.1) if $a \in f(\phi)$, then outcome a is implemented; STOP
  - (0.2.2) if $a \notin f(\phi)$ but i is not the agent $j(0)$ prescribed in $\sigma^+(\theta, \phi; a)$, then outcome a is implemented; STOP

- (0.2.3) if $a \notin f(\phi)$ and $i = j(0)$, then go to Stage 1.

- (0.3) If neither (0.1) nor (0.2) applies, then the agent with the highest integer n_i^0 is allowed to choose an outcome from B. Ties are broken by selecting from the agents announcing the highest integer the one with the smallest i. STOP

(b) Subsequent Stages (Stages $k = 1, 2, \cdots, l$): Each agent $i = 1, 2, \cdots, I$ can either raise a "flag" or announce a non-negative integer n_i^k. Again there are three possibilities to consider:

- (k.1) If $I - 1$ or more flags are raised, then the agent $j(k-1)$ prescribed in $\sigma^+(\theta, \phi; a)$ is allowed to choose an outcome from B. STOP
- (k.2) If $I - 1$ or more agents announce zero, and
  - (k.2.1) if the agent $j(k)$ prescribed in $\sigma^+(\theta, \phi; a)$ is one of those who announce zero, then implement outcome a_k from sequence $\sigma^+(\theta, \phi; a)$; STOP
  - (k.2.2) if $j(k)$ does not announce zero, then

    if $k < l$, go to Stage $k + 1$; and

    if $k = l$, then implement outcome a_{l+1} from sequence $\sigma^+(\theta, \phi; a)$. STOP
- (k.3) If neither (k.1) nor (k.2) applies, then the agent who announces the highest integer n_i^k is allowed to choose an outcome from B. (In this comparison, raising a flag counts as -1, say.) Ties are broken as in (0.3). STOP

Now, we claim that for each true state θ^*, and for each $a \in f(\theta^*)$, it is an SPNE in g where for all i, agent i announces (θ^*, a, n_i^0) at Stage 0 and would announce zero at Stage $k \in \{1, 2, \cdots, l\}$ if the latter stages were ever reached. The game stops at Stage 0 in this SPNE with a bein the equilibrium outcome. Moreover, these are the only possible SPNE's. Hence g fully implements f in SPNE.

To prove this assertion, we first show that

Lemma SP. In each Stage $k \in \{1, 2, \cdots, l\}$ of g, there can be two possible SPNE's only, referred to as F_k and Z_k respectively, where in equilibrium F_k, all agents would raise a flag, and in equilibrium Z_k, all agents would announce zero.

To prove Lemma SP, note first that (k.3) can never be an equilibrium at Stage k. Thus apart from F_k and Z_k, there are two possible equilibria to consider: where exactly $I-1$ agents raise a flag, and where exactly $I-1$ agents announce zero. Next recall that these stages can be reached only if at Stage 0, $I-1$ agents agree on (θ, a) and the remaining agent $j(0)$ announces ϕ with $a \in f(\theta) \setminus f(\phi)$.

To see that it is not an equilibrium in Stage k where exactly $I-1$ agents raise a flag, note that it is some outcome contained in $M^{j(k-1)}(\theta^*) \subset B$ that would be implemented in this supposed equilibrium. Since $I-1 \geq 2$ there must be some agent $i \neq j(k-1)$ currently raising a flag who can deviate and announce a sufficiently high n_i^k to implement an outcome in $M^i(\theta^*) \subset B$ instead.

To see that it is not an equilibrium in Stage k where exactly $I-1$ agents announce zero, we take cases. At first, if $j(k)$ is one of those announcing zero, then $a_k \in Q^+$ would be implemented, but since $M^{j(k)}(\theta^*) \bigcap Q^+ = \emptyset$, a_k is not a state-θ^* top-ranked choice for $j(k)$. Thus $j(k)$ would be better off announcing a high $n_{j(k)}^k$ instead, to implement some top-ranked outcome in B. Next, if $j(k)$ is not one of those announcing zero, then the game would move on to Stage $k+1$, where the implemented outcome cannot be top-ranked for each and every agent announcing zero in Stage k. Thus some agent j currently announcing zero can deviate and announce a high n_j^k to implement his top-ranked outcome in B. To sum up, it is not an equilibrium in Stage k where exactly $I-1$ agents announce zero.

Now, it is easy to see that at Stage $k \in \{1, 2, \cdots, l\}$ of g, F_k is *always* an equilibrium: unilateral deviations would not alter anything, some top-ranked outcome for $j(k-1)$ would still be implemented.

At Stage $k \in \{1, 2, \cdots, l\}$ of g, is Z_k an equilibrium in true state θ^*? Yes, if the θ announced by $I-1$ agents in Stage 0 is exactly the true state θ^*, and if at Stage k all agents are expecting Z_{k+1} in Stage $k+1$: agent $j(k)$ would get a_k in equilibrium Z_k, and by deviating he would expect to get a_{k+1}, while $a_k R^{j(k)}(\theta) a_{k+1}$. Note that any other agent cannot deviate from Z_k to implement an outcome differing from a_k. Thus Z_k is an equilibrium at Stage k if $\theta = \theta^*$ and if Z_{k+1} would prevail when Stage $k+1$ were reached.

Now we ask the crucial question: Is Z_l an equilibrium at Stage l? Recall that a_l would be implemented in Z_l while a_{l+1} would be implemented if $j(l)$ deviates and chooses to not announce zero. If $\theta = \theta^* \neq \phi$, then $j(l)$ would announce zero, since $a_l R^{j(l)}(\theta) a_{l+1}$; but if $\phi = \theta^* \neq \theta$, then $j(l)$ would deviate to implement a_{l+1} because $a_{l+1} P^{j(l)}(\phi) a_k$! We conclude that Z_l is an equilibrium at Stage l when the $I-1$ agents announcing θ are telling the truth.

Let us sum up what we have learned so far. Suppose that following the agents' announcements at Stage 0, it turns out that (0.2.3) applies. Then one equilibrium outcome of the subgame starting at Stage 1 is that agent $j(0)$ chooses his top-ranked outcome from B (equilibrium F_1). This will be the unique equilibrium outcome whenever $\phi = \theta^* \neq \theta$. However if $\theta = \theta^* = \phi$, then there is a second equilibrium (equilibrium Z_1), with outcome a_1.

Now, summarizing the above discussions, we can examine Stage 0. The integer game (0.3) clearly has no equilibria. Also, it cannot be an equilibrium where (0.2) applies, with exactly $I-1$ agents agreeing upon (θ, a). (Here the reasoning is the same as that given above to show that at Stage $k > 0$ it is not an equilibrium for exactly $I-1$ agents to announce zero.) Thus in equilibrium, all agents must announce the same (θ, a) with $a \in f(\theta)$. Now, this cannot be an equilibrium either if $\theta \neq \theta^*$ and $a \notin f(\theta^*)$. Why not? Because the agent $j(0)$ prescribed in $\sigma^+(\theta, \theta^*; a)$ can do strictly better by announcing $\phi = \theta^*$ in order to be able to choose his state-θ^* top-ranked outcomes from B at Stage 1 (where F_1 is the unique equilibrium).

We conclude that in equilibrium at Stage 0, all agents must announce the same (θ, a) with $a \in f(\theta) \bigcap f(\theta^*)$, so that a is implemented. Thus the set of state-θ^* SPNE outcomes under g is a subset of $f(\theta^*)$.

Finally, we show that the set of state-θ^* SPNE outcomes under g contains $f(\theta^*)$. Indeed, for each $a \in f(\theta^*)$, consider letting agent i announce (θ^*, a, n_i^0) at Stage 0 and zero at Stage $k > 0$. Nobody would unilaterally deviate from this strategy profile,[22] which has a as the SPNE outcome, and thus the set of state-θ^* SPNE outcomes under g does contain $f(\theta^*)$.

This finishes the proof.

70. Part 4. **Bayesian Implementation**.

71. We shall focus only on the implementability of an SCF (i.e. single-valued SCC). Assume that there are I agents, and agent i's information is his own type $\theta_i \in \Theta^i$, and agent i is endowed with a family of conditional distributions

$$G^i(\theta_{-i}|\theta_i), \forall \theta_i \in \Theta^i, \ \forall \theta_{-i} \in \Theta^{-i} \equiv \Pi_{j \neq i} \Theta^j,$$

and a state-dependent utility function

$$U^i(a, \theta_i), \ \forall a \in A, \ \forall \theta_i \in \Theta^i.$$

We can derive an interim utility function from U^i via

$$V^i(f(\cdot), \theta_i) \equiv \int_{\Theta^{-i}} U^i(f(\theta), \theta_i) dG^i(\theta_{-i}|\theta_i).$$

A mechanism or game form (g, M) is defined as before, with M^i being the message space for agent i, and $M \equiv \Pi_i M^i$. A pure strategy for agent i under this game form is a function $s^i : \Theta^i \to M^i$, and let S^i denote agent i's strategy space (which consists of all his pure strategies) and $S \equiv \Pi_i S^i$.

72. **Theorem B1.** (**Revelation Principle.**) Suppose that (g, M) implements the SCF f in Bayesian equilibrium, in the sense that there exists $s = (s_1, s_2, \cdots, s_I)$ such that for all i and for all $\theta \in \Theta$, and for all $s_i' \in S^i$,

[22] Recall that Z_k is indeed a Stage-k equilibrium when $I-1$ agents are telling the true state θ^*, but then $j(k)$ would prefer sticking to Z_k with outcome a_k than moving on to Stage $k+1$ with outcome a_{k+1}.

$$\int_{\Theta^{-i}} U^i(g(s(\theta)), \theta_i) dG^i(\theta_{-i}|\theta_i) \geq \int_{\Theta^{-i}} U^i(g(s_{-i}(\theta_{-i}), s_i'(\theta_i)), \theta_i) dG^i(\theta_{-i}|\theta_i),$$

and

$$g(s(\theta)) = f(\theta),$$

then the direct game form (f, Θ) also truthfully implements f in Bayesian equilibrium.

Proof. Suppose not. Then in some state $\theta \in \Theta$, some agent i would strictly prefer reporting θ_i' while expecting his rival agents to truthfully report θ_{-i}, implying that

$$\int_{\Theta^{-i}} U^i(f(\theta_{-i}, \theta_i'), \theta_i) dG^i(\theta_{-i}|\theta_i) > \int_{\Theta^{-i}} U^i(f(\theta_{-i}, \theta_i), \theta_i) dG^i(\theta_{-i}|\theta_i).$$

Since

$$g(s(\cdot)) = f(\cdot),$$

this implies that

$$\int_{\Theta^{-i}} U^i(g(s_{-i}(\theta_{-i}), s_i(\theta_i')), \theta_i) dG^i(\theta_{-i}|\theta_i)$$

$$> \int_{\Theta^{-i}} U^i(g(s_{-i}(\theta_{-i}), s_i(\theta_i)), \theta_i) dG^i(\theta_{-i}|\theta_i) = \int_{\Theta^{-i}} U^i(g(s(\theta)), \theta_i) dG^i(\theta_{-i}|\theta_i),$$

so that agent i in playing the game (g, M) would strictly prefer the message $s_i(\theta_i')$ to $s_i(\theta_i)$ when his type is θ_i, a contradiction to the assumption that (g, M) implements f in Bayesian equilibrium.

73. When the direct game form (f, Θ) truthfully implements f, we say that f is *incentive compatible* in the sense that in the truthful equilibrium under (f, Θ), for all i, when expecting all agents $j \neq i$ to always report their true types, agent i would also find it optimal to report his true type.

74. We shall denote by $\alpha_i : \Theta^i \to \Theta^i$ a typical pure strategy for i under a direct game form, and refer to it as a *deception*. If $\alpha_i(\theta_i) = \theta_i$ for all $\theta_i \in \Theta^i$, then α_i is referred to as the identity deception.

Thus when agent i's true type is θ_i, agent i reports that his type is $\alpha_i(\theta_i)$ when he adopts the pure strategy α_i. Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_I)$ denote a joint deception for the I agents. If $\alpha(\theta)$ is one equilibrium where $f(\alpha(\theta)) \neq f(\theta)$, then α is an unwanted equilibrium.

75. Again, truthful implementation is weak. Even if (g, M) implements f, its equivalent direct game form (f, Θ) may not, because there may be unwanted Bayesian equilibria arising under (f, Θ).

The following is an example. Suppose that $I = 2$, and each agent i has two possible types θ_{i1} and θ_{i2}, so that Θ has four elements, which we assume are equally likely. Suppose that $A = \{a, b, c\}$, and let $U_{in} : A \to \Re$ be agent i's utility function when his type is θ_{in}, where we assume that

	U_{11}	U_{12}	U_{21}	U_{22}
a	2	0	2	2
b	1	4	1	1
c	0	9	0	-8

Note that in the current example, each agent has 4 feasible pure strategies (or deceptions) under a direct game form: *to always tell the truth, to always lie, to always report type 1*, and *to always report type 2*.

The following matrix denotes f, and it is truthfully implementable.

	θ_{21}	θ_{22}
θ_{11}	a	b
θ_{12}	c	b

Indeed, each agent has no reason to lie, if expecting the other agent to always tell the truth.[23] However, this direct game form has another Bayesian equilibrium, where agent 1 always tells the truth, but agent 2 always reports type θ_{22}, so that the equilibrium outcome is always b.[24] Note that

[23] Assuming that the rival would always tell the true, the incentive compatibility condition for θ_{11} is satisfied, as $\frac{1}{2}(2+1) > \frac{1}{2}(0+1)$. Similarly, we have $\frac{1}{2}(9+4) > \frac{1}{2}(0+4)$ for θ_{12}, $\frac{1}{2}(0+2) \geq 1$ for θ_{21}, and $1 > \frac{1}{2}(-8+2)$ for θ_{22}.

[24] Again, multiple equilibria arise because θ_{21} feels indifferent about getting b for sure or getting a or c with equal probability.

$$b \notin f(\theta_{11}, \theta_{21}),\ b \notin f(\theta_{12}, \theta_{21}),$$

and hence this *pooling equilibrium* is an unwanted equilibrium.

76. Note that in a Bayesian equilibrium, by definition agent 1 can correctly expect agent 2's equilibrium strategy. Thus we can give agent 1 one more strategy ϕ (to send a message called "objection") to eliminate the unwanted equilibrium in the preceding example. This procedure is called *selective elimination*, and is due to Mookherjee and Reichelstein (1990).[25] The augmented game form becomes

	θ_{21}	θ_{22}
θ_{11}	a	b
θ_{12}	c	b
θ_{11}, ϕ	b	a
θ_{12}, ϕ	b	c

where note that when agent 1 says "objection" then the outcome is switched between θ_{21} and θ_{22}.

This modification of the original direct game form attains three goals: it removes the unwanted equilibrium; it retains the desired equilibrium, and it does not introduce any new unwanted equilibria. We leave the reader to verify that this new game form does implement f.[26]

[25] Mookherjee, D., and S. Reichelstein, 1990, Implementation via Augmented Revelation Mechanisms, *Review of Economic Studies*, 57, 453-475.

[26] We proceed as follows.

- At first, truth-telling remains an equilibrium: since agent 1 considers θ_{21} and θ_{22} equally likely, he can safely disregard the last two rows when expecting agent 2 to always tell the truth; and expecting agent 1 to always tell the truth, we have already verified in footnote 23 that truth-telling is optimal for both θ_{21} and θ_{22}.
- Now, we show that in equilibrium agent 2 would not always claim to be θ_{22}. If this were an equilibrium, then θ_{11}'s best response is to reveal his own identity and to say "objection," so that a would be implemented; and θ_{12}'s best response is to reveal his own identity and to say "objection," so that c would be implemented. But then, θ_{22} would be better off claiming to be θ_{21}, a contradiction.

77. The preceding example shows that if we have $f(\alpha(\theta)) \neq f(\theta)$, then we must make sure that α is not a Bayesian equilibrium in an augmented version of the direct game form (f, Θ). We need to have, in this situation, some agent i who can adopt a new strategy ϕ_i (as agent 1's saying "objection" in the preceding example) and when the latter happens a new social choice function y will be implemented, which encourages agent i to use ϕ when the other agents use α_{-i}. Moreover, we must make sure that such a modification does not affect the equilibrium outcome whenever all agents are telling the truth. The latter desired property is referred to as *Bayesian monotonicity.*

78. **Definition B1.** The SCF $f : \Theta \to A$ is Bayesian monotonic if for any deception α such that $f(\alpha(\theta)) \neq f(\theta)$ at some $\theta \in \Theta$, there exists some agent i, his type θ_i, and another SCF y depending only on θ_{-i} such that

$$V^i(f(\alpha(\theta)), \theta_i) < V^i(y(\alpha_{-i}(\theta_{-i}), \phi_i)), \theta_i),$$

and yet, for all $\theta_i' \in \Theta^i$,

$$V^i(f(\theta'), \theta_i') \geq V^i(y(\theta_{-i}', \phi_i), \theta_i').$$

Note that Bayesian monotonicity extends the definition of monotonicity defined previously in section 17.

79. **Theorem B2.** Suppose that f is Bayesian implementable, then f is both *incentive compatible* and *Bayesian monotonic.*

- By the same reasoning, in equilibrium agent 2 would not always claim to be θ_{21}. Indeed, if this were an equilibrium, then agent 1's best response is to always tell the truth, but then θ_{22} would be better off deviating and telling the truth instead.
- Now, would there be an equilibrium where agent 2 always tells a lie? In such an equilibrium, for both $n = 1, 2$, θ_{1n}'s equilibrium message would either be θ_{1n} simply, or θ_{1n} and say "objection, but for θ_{22} to pretend to be θ_{21}, θ_{12} must choose to say "objection." Now, since θ_{21} prefers b to c, for θ_{21} to pretend to be θ_{22}, it must be that θ_{11} would choose to say "objection " as well. Thus we do have an equilibrium where agent 2 always tells a lie and agent 1 always says "objection."

To summarize, we obtain two equilibria sharing the same outcome. Thus there are no unwanted equilibria.

80. For all i, let Y^{-i} denote the set of SCF $y(\theta_{-i}, \phi_i)$. The following result holds in a very special environment.

Theorem B3. Suppose that the SCF f is both *incentive compatible* and *Bayesian monotonic*, and that $I \geq 3$. Suppose that for each $i \in \{1, 2, \cdots, I\}$, there exists a unique $b_i \in A$ such that $L^i(b_i, \theta_i) = A$ for all $\theta_i \in \Theta^i$, where $b_i \neq b_j$ for all $i \neq j$, $i, j \in \{1, 2, \cdots, I\}$. Suppose also that agents have common prior distribution $G(\theta) > 0$ for all $\theta \in \Theta$, with Θ being a finite set. Then f is Bayesian implementable.

Proof. Consider the game form (g, M) with $M^i = \Theta^i \times \{Y^{-i} \bigcup \{0, 1, 2, \cdots\}\}$, and with g being such that

(a) $g(m) = f(\theta)$ if for all i, $m^i = (\theta_i, 0)$;

(b) if for some i, $m^j = (\theta_j, 0)$ for all $j \neq i$, and $m^i = (\theta_i, k_i)$ with $k_i > 0$ or $m^i = (\theta_i, y(\cdot, \phi_i))$, then

$$g(m) = \begin{cases} f(\theta), & \text{if } m^i = (\theta_i, k_i) \text{ with } k_i > 0 \\ & \text{or if for some } \theta_i', \ V^i(f(\theta'), \theta_i') < V^i(y(\theta_{-i}', \phi_i), \theta_i'); \\ y(\cdot, \phi_i), & \text{if } V^i(f(\theta'), \theta_i') \geq V^i(y(\theta_{-i}', \phi_i), \theta_i') \text{ for all } \theta_i' \in \Theta^i. \end{cases}$$

(c) $g(m) = b^j$ for any other joint message m, where agent j has announced the highest integer and ties are broken by selecting the agent with the smallest i from among those agents announcing the highest integer.

We shall show that an equilibrium must take the form of $m^i = (\theta_i, 0)$ for all i. When all agents are telling the truth, unilateral deviation is not worthwhile, but when all agents are not telling the truth, then by Bayesian monotonicity some test agent i can adopt ϕ_i and implement $y(\cdot, \phi_i)$ and become better off.

Now we give details.

At first, we claim that in any true state θ, $m^i = (\theta_i, 0)$ for all i constitutes an equilibrium. Note that if all agents other than i are telling the truth, then $f(\theta_i, \theta_{-i})$ would be implemented according to g, unless agent i announces some $(\theta_i, y(\cdot, \phi_i))$ with $V^i(f(\theta'), \theta_i') \geq V^i(y(\theta_{-i}', \phi_i), \theta_i')$ for all $\theta_i' \in \Theta^i$. Since f is incentive compatible, given agent i's true type θ_i, reporting $\hat{\theta}_i = \theta_i$ is optimal for agent i if he expects $f(\hat{\theta}_i, \theta_{-i})$ to be implemented. On the other hand, announcing some $(\theta_i, y(\cdot, \phi_i))$ is obviously sub-optimal for agent i given that $V^i(f(\theta'), \theta_i') \geq V^i(y(\theta_{-i}', \phi_i), \theta_i')$ for all $\theta_i' \in \Theta^i$ and given that all agents other than i are using the (truth-telling) identity deception. Thus no unilateral deviation can make agent i better off, and this proves that $m^i = (\theta_i, 0)$ for all i constitutes an equilibrium, which implements $f(\theta)$ as the desired outcome.

Next, observe that $m^i = (\alpha_i(\theta_i), 0)$ for all i does not constitute an equilibrium in true state θ if $f(\alpha(\theta)) \neq f(\theta)$. Indeed, in this case by Bayesian monotonicity there exists agent i, some $\theta_i \in \Theta^i$, and some SCF $y(\cdot, \phi_i)$ such that agent i given his type θ_i would deviate and send the message $m^{i\prime} = (\theta_i, y(\cdot, \phi_i))$ and become better off, given that all other agents are reporting $\alpha_{-i}(\theta_{-i})$.

Finally, there does not exist an equilibrium where some agent i would announce a strictly positive integer, or announce some SCF from Y^{-i} at some $\theta_i \in \Theta^i$. This happens because Θ is a finite set with $G(\cdot) > 0$ on Θ (so that no single type of agent i can make those deviations and be ignored by his rival agents) and because the top ranked elements in A by the I agents are all distinct and state-invariant. If there were such an equilibrium, then each agent i alone could and would deviate and announce an integer k_i exceeding the integers announced by his $I-1$ rival agents, in an attempt to implement b_i. This is a contradiction.

81. We have so far assumed that there are a fixed number I of agents that would play the game form. In some cases, adding one more uninformed agent can turn a non-implementable SCF into one that is implementable.[27] Consider the following example.

There is a principal, and two agents, called γ and β. The principal would design the game form. Each of the two agents can be of either type 1 (with probability q) or type 2, and their types are independently

[27]One should compare this result to the fact that in Nash and SP Nash implementation scenarios, f is more likely to be implementable with $I \geq 3$ than with $I = 2$.

distributed. Thus there are four possible preference states:

$$\Theta = \{\gamma_1\beta_1, \gamma_1\beta_2, \gamma_2\beta_1, \gamma_2\beta_2\}.$$

There are three feasible choices: $A = \{a, b, c\}$. One interpretation is that the principal can undertake project a or project b or firing both agents (project c).

Assume that each of the three people would get payoff 1 if his favorite choice is implemented; payoff 0 if the choice he hates most is implemented, and payoff $v \in (0, 1)$ if the remaining choice is implemented, where

$$v > \max(q, 1 - q).$$

The table below summarizes the two agents' common preferences in each preference state:

$\gamma_1\beta_1$	$\gamma_1\beta_2$	$\gamma_2\beta_1$	$\gamma_2\beta_2$
a	a	b	b
b	c	c	a
c	b	a	c

Table A.

The table below summarizes the principal's preferences in each preference state:

$\gamma_1\beta_1$	$\gamma_1\beta_2$	$\gamma_2\beta_1$	$\gamma_2\beta_2$
b	c	c	a
c	b	a	c
a	a	b	b

Table P.

Thus the principal has exactly the opposite preferences over $\{a, b\}$ to the two agents. The SCF f that the principal would like to implement is the 2nd row in Table P, which can be truthfully implemented by the following direct game form:[28]

	β_1	β_2
γ_1	b	c
γ_2	c	a

However, this game form has an unwanted equilibrium, where both agents always lie.[29] In fact, in this example, f is *not* Bayesian implementable even if the central planner is allowed to choose any *two-agent* indirect game forms.[30]

Palfrey (1990) shows that adding an uninformed player can help.[31] In particular, if the principal joins the game with a message space { truth, lie }, then the three-player simultaneous game form represented by the following two matrices

	β_1	β_2	
γ_1	b	c	truth
γ_2	c	a	

[28]For γ_1, truth-telling generates v, and pretending to be γ_2 would generate $q \cdot 0 + (1 - q) \cdot 1 < v$. Similarly, we have $v > q \cdot 1 + (1 - q) \cdot 0$ for γ_2, and $v > q \cdot 0 + (1 - q) \cdot 0$ for β_1 and β_2. Thus truth-telling is a Bayesian equilibrium.

[29]Expecting the rival to always lie, γ_1 would get $q \cdot 0 + (1 - q) \cdot 0$ by telling the truth, or $q \cdot 1 + (1 - q) \cdot v$ by pretending to be γ_2. To see this, note that, for example, when γ_1 is telling the truth, with probability q agent 2 is of type β_1, who is pretending to be β_2, so that $f(\gamma_1, \beta_2) = c$ would be implemented, generating 0 for γ_1 in state $\gamma_1\beta_1$; and with probability $1 - q$ agent 2 is of type β_2, who is pretending to be β_1, so that $f(\gamma_1, \beta_1) = b$ would be implemented, generating 0 for γ_1 in state $\gamma_1\beta_2$. Similarly, we have $q \cdot 0 + (1 - q) \cdot 0 < q \cdot v + (1 - q) \cdot 1$ for γ_2; $q \cdot 0 + (1 - q) \cdot 1 < q \cdot 1 + (1 - q) \cdot v$ for β_1; and $q \cdot 1 + (1 - q) \cdot 0 < q \cdot v + (1 - q) \cdot 1$ for β_2. Thus there is a Bayesian equilibrium where both agents always lie about their types.

[30]This happens because the SCF violates Bayesian monotonicity. Define $\theta \equiv \gamma_1\beta_1$ and $\theta' \equiv \gamma_2\beta_2$. In the unwanted equilibrium, in state $\theta = \gamma_1\beta_1$, $f(\alpha(\theta)) = f(\theta') = a \neq b = f(\theta)$; and in state $\theta' = \gamma_2\beta_2$, $f(\alpha(\theta')) = f(\theta) = b \neq a = f(\theta')$. Note that a is uniquely top-ranked by both agents in state θ and b is uniquely top-ranked by both agents in state θ'. Thus it is impossible to find in state θ some agent i and some SCF $y(\cdot, \phi_i)$ such that agent i would prefer $y(\alpha_{-i}(\theta_{-i}), \phi_i)$ to $f(\alpha(\theta))$.

[31]Palfrey, T., 1990, Implementation in Bayesian Equilibrium: The Multiple Equilibrium Problem in Mechanism Design, in J.-J. Laffont (Ed.) *Advances in Economic Theory*, Cambridge: Cambridge University Press.

	β_1	β_2
γ_1	a	c
γ_2	c	b

lie

has two equilibria, both of which produce the outcome specified by f. In one equilibrium, the two agents always tell the truth and the principal would say "truth," and in the other equilibrium, the two agents always lie and the principal would say "lie," and in the latter equilibrium the interpretation of the two agents' messages is reversed. Thus the unwanted bad equilibrium is not actually eliminated, but it is converted into a good equilibrium. Verifying the two equilibria is left as an exercise.

Adding an uninformed player can help because, once again, by definition of Bayesian equilibrium, agents' strategies are correctly expected by the newly added player in equilibrium.

82. In many cases, unlike the preceding example, people are faced with the allocation of a transferable money good in fixed supply. The presence of such a money good may sometimes make an originally non-implementable SCF implementable.

83. **Definition B2.** Let $G^j(\theta_i)$ be the beliefs that agent j has about agent i's type. Let $G^j_\alpha(\theta_i)$ be the j-believed probability distribution of reported types by agent i under the joint deception α. Then the environment satisfies *no consistent deceptions* (NCD) if priors are objective and for all i, j, $G^j(\theta_i) = G^j_\alpha(\theta_i)$ for all $\theta_i \in \Theta^i$ implies that agent i is telling the truth under α.

    In words, under NCD, truthful reporting is the *only* strategy that an agent has which produces a probability distribution of his reported types that is exactly the same as the distribution of his types that the other agents believe.

    Theorem B4. Suppose that types are independent, NCD is satisfied, and there is a money good. Suppose that agent i's utility function is $v(a, \theta_i) + w_i$, for all $a \in A$ and $\theta_i \in \Theta^i$, and where w_i is the amount of money good possessed by agent i. Then an SCF f is Bayesian implementable if and only if f is incentive compatible.

Proof. The "only if" part follows from Theorem B2. To prove the "if" part, we construct a game form (g, M). Define

$$W_i \equiv \{w : \Theta^{-i} \to \Re^I | \ \ \textstyle\sum_{j=1}^{I} w_j(\theta_{-i}) = 0, \ \forall \theta_{-i} \in \Theta^{-i}$$

$$\textstyle\sum_{\theta_{-i}} G^i(\theta_{-i}) w_i(\theta_{-i}) < 0\}.$$

Let $M^i = \Theta^i \times (W_i \bigcup \{0\})$ for all i, and for all $m \in M$,

$$g(m) = \begin{cases} f(\theta), & \text{if } \ m^i = (\theta_i, 0), \ \forall i; \\ \\ f(\theta) \text{ plus } \ w(\theta_{-i}), & \text{if otherwise, where } i \text{ is the lowest-indexed agent} \\ & \text{not announcing zero.} \end{cases}$$

Now, in true state θ, if agent i expects all his rival agents would truthfully announce θ_{-i} and also announce zero, then by the fact that f is incentive compatible agent i would also truthfully report his type θ_i and announce zero; reporting some $w \in W_i$ rather than zero is not worthwhile because $\sum_{\theta_{-i}} G^i(\theta_{-i}) w_i(\theta_{-i}) < 0$.

It is not an equilibrium in true state θ for all agents to report their types truthfully and for some agents to not report zero, because the agent i determining $w(\cdot)$ is better off replacing $w_i(\cdot)$ by $kw_i(\cdot)$ with some $k \in (0, 1)$.

It remains to show that there can be no equilibrium with a non-identity deception α. At first, it cannot be an equilibrium with everyone reporting $(\alpha_i(\theta_i), 0)$, since by NCD, there exists some agent i who can replace his equilibrium message $(\alpha_i(\theta_i), 0)$ by the deviation message $(\alpha_i(\theta_i), w)$, where $w \in W_i$ is such that

$$\sum_{\theta_{-i}} G^i(\theta_{-i}) w_i(\theta_{-i}) < 0 < \sum_{\theta_{-i}} G^i(\alpha(\theta_{-i})) w_i(\alpha(\theta_{-i}));$$

the latter inequality can hold because of NCD.

Finally, there cannot be an equilibrium with a non-identity deception and with a subset of agents reporting various $w(\cdot) \neq 0$. In the latter supposed equilibrium, let i be the one determining w, and there are three possibilities: either

$$\sum_{\theta_{-i}} G^i(\alpha(\theta_{-i}))w_i(\alpha(\theta_{-i})) < 0,$$

so that agent i would deviate and replace w by kw, where $k \in (0,1)$; or either

$$\sum_{\theta_{-i}} G^i(\alpha(\theta_{-i}))w_i(\alpha(\theta_{-i})) > 0,$$

so that agent i would deviate and replace w by kw, where $k > 1$; or

$$\sum_{\theta_{-i}} G^i(\alpha(\theta_{-i}))w_i(\alpha(\theta_{-i})) = 0,$$

and in the latter case by NCD there exists some $\hat{w} \in W_i$ such that

$$\sum_{\theta_{-i}} G^i(\theta_{-i})\hat{w}_i(\theta_{-i}) < 0 = \sum_{\theta_{-i}} G^i(\alpha(\theta_{-i}))w_i(\alpha(\theta_{-i}))$$

$$< \sum_{\theta_{-i}} G^i(\alpha(\theta_{-i}))\hat{w}_i(\alpha(\theta_{-i})),$$

and hence agent i would deviate and replace w by $\hat{w}$ instead.

To sum up, f is implemented by the game form (g, M).

84. As an application of Theorem B4, recall the example in section 81. Suppose that $q > \frac{1}{2}$.[32] Recall that without adding a new uninformed player, the principal's SCF f is not Bayesian implementable. Let f_L denote the choice function implemented in the bad equilibrium where both agents always lie.

Now, assume that there exists a money good. In addition to f, consider the money transfer rule w: if β reports type 1, then γ must pay β one dollar, and if β reports type 2, then β must pay γ an amount $\frac{q}{1-q} - \epsilon$. For $\epsilon > 0$ small enough, this rule w makes γ better off than in f_L if β always lie, and it makes γ worse off than in f if β always tells the

[32] Note that NCD is satisfied in this case.

truth,[33] and this conclusion is true regardless of γ's type, following the assumption of independent types.

Allowing γ alone to choose between a class of such money transfer rules w and zero (no such transfer rule) would thus be useful for removing the unwanted equilibrium. Expecting β to always tell the truth, γ would rather choose zero; and in expecting β to always lie, γ would choose some money transfer rule to benefit himself, which penalizes β severely via $w_\gamma + w_\beta = 0$. In equilibrium, as Theorem B4 shows, β would always tell the truth, and γ would choose zero.

85. Part 5. **Durability of Bayesian Mechanisms**.

86. We have thus far considered the case where a game form is selected for a group of I agents by a central planner. In many applications, however, it is the I agents bargaining over the design of a game form, at or before the time that the I agents receive private information about the true state θ. Holmström and Myerson (1983) consider the latter scenario,[34]

87. In a standard model with asymmetric information, agent $i \in \{1, 2, \cdots, I\}$ may hold prior beliefs $p_i(\theta)$ over $\theta \in \Theta$ at the ex-ante stage, and may then form the posterior beliefs $p_i(\theta_{-i}|\theta_i)$ after learning about his own type θ_i at the interim stage, and then observe the realized true state θ at the ex-post stage.[35]

Holmström and Myerson call an SCC $f : \Theta \to A$ a *decisoin rule*, and they ask which decision rules should be considered *efficient*.

In a model with symmetric information, where all agents are endowed with full knowledge about the realized θ, efficiency can be defined as Pareto efficiency: f is efficient if no other decision rule g Pareto dominates f in any state θ.

With information asymmetry, one approach, which Holmström and Myerson refer to as *classical*, is to assume that there are no incen-

[33]Note that $(1-q) \cdot (-1) + q \cdot (\frac{q}{1-q} - \epsilon) > 0$ because $q^2 > (1-q)^2$ and $\epsilon > 0$; and that $q \cdot (-1) + (1-q) \cdot (\frac{q}{1-q} - \epsilon) < 0$ because $\epsilon > 0$.

[34]Holmström, B., and R. Myerson, 1983, Efficient and Durable Decision Rules With Incomplete Information, *Econometrica*, 51, 1799-1819.

[35]It is assumed that p_i and p_j are equivalent probability measures, in the sense that they assign the same zero-probability events.

tive problems involved in eliciting the necessary information $\{\theta_i; i = 1, 2, \cdots, I\}$ from individuals, when implementing a decision rule f. Let Δ denote the set of classically feasible decision rules.

In general, however, a decision rule $f \in \Delta$ may not really be implementable, if it depends on information that individuals hold privately and that they do not want to reveal. Thus we say that f is *incentive feasible* or *incentive compatible* if and only if

$$\text{(IC)} \quad \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(f(\theta), \theta) \geq \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(f(\theta_{-i}, \hat{\theta}_i), \theta), \ \forall i, \ \forall \theta_i, \hat{\theta}_i \in \Theta_i.$$

Let

$$\Delta^* \equiv \{f \in \Delta : \ f \text{ satisfies (IC)}\}.$$

Now, define

$$U_i(f) \equiv \sum_{\theta} p_i(\theta) u_i(f(\theta), \theta),$$

$$U_i(f|\theta_i) \equiv \sum_{\theta_{-i}} p_i(\theta_{-i}|\theta_i) u_i(f(\theta), \theta),$$

and

$$U_i(f|\theta) \equiv u_i(f(\theta), \theta).$$

Now, we say that a decision rule g *ex-ante dominates* f if and only if

$$U_i(g) \geq U_i(f), \ \forall i,$$

with at least one strict inequality; g *interim dominates* f if and only if

$$U_i(g|\theta_i) \geq U_i(f|\theta_i), \ \forall i, \ \forall \theta_i \in \Theta_i,$$

with at least one strict inequality; and g *ex-post dominates* f if and only if

$$U_i(g|\theta) \geq U_i(f|\theta), \ \forall i, \ \forall \theta \in \Theta,$$

with at least one strict inequality.

Notice that in the interim and ex post cases, domination requires (weakly) increasing expected utility for all possible types, not just for those in the actual information state of the economy. This requirement is necessary because a welfare economist, as an outsider, could not apply any concept of domination that depended on the individuals' actual private information. Therefore, we must compare I utility measures in the ex ante case, $\sum_{i=1}^{I} |\Theta_i|$ utility measures in the interim case, and $I \cdot |\Theta|$ utility measures in the ex post case, where recall that $|C|$ denotes the number of elements contained in the set C.

The three notions of domination and two notions of feasibility (Δ and Δ^*) together generate six potential concepts of efficiency. Let Δ_A (respectively, Δ_I, and Δ_P) denote the set of *ex-ante* (respectively, *interim*, and *ex-post*) classically efficient decision rules. Correspondingly, let Δ_A^* (respectively, Δ_I^*, and Δ_P^*) denote the set of *ex-ante* (respectively, *interim*, and *ex-post*) incentive efficient decision rules. It is easy to see that

$$\Delta_A \subset \Delta_I \subset \Delta_P, \quad \Delta_A^* \subset \Delta_I^* \subset \Delta_P^*.$$

Moreover, because

$$\Delta^* \subset \Delta,$$

we also have

$$\Delta_A \bigcap \Delta^* \subset \Delta_A^*; \ \Delta_I \bigcap \Delta^* \subset \Delta_I^*; \ \Delta_P \bigcap \Delta^* \subset \Delta_P^*.$$

Holmström and Myerson emphasize that it may happen that

$$\Delta_P \bigcap \Delta^* = \emptyset,$$

so that no classically efficient decision rules (in any sense) can be incentive compatible. The following is an example.

There are two individuals, each individual has two equally likely types, $\Theta_1 = \{1a, 1b\}$ and $\Theta_2 = \{2a, 2b\}$, and each individual's type is stochastically independent of the other's. Suppose that $A = \{\alpha, \beta\}$, and the two individuals' state-dependent payoffs $(u_1, u_2) : A \to \Re^2$ are as follows.

	$(1a, 2a)$	$(1a, 2b)$	$(1b, 2a)$	$(1b, 2b)$
α	(6,0)	(0,0)	(2,2)	(0,0)
β	(0,6)	(2,2)	(0,0)	(2,2)

If f attains ex-post classical efficiency, then we must have

$$f(1a, 2b) = \beta = f(1b, 2b), \; f(1b, 2a) = \alpha.$$

If type $1a$ pretended to be $1b$, assuming that agent 2 is always truthful, then he expects to implement α and β with equal probability, so that his expected utility from lying is

$$\text{prob.}(1a, 2a) \cdot 6 + \text{prob.}(1a, 2b) \cdot 2 = 4.$$

Thus for f to be incentive compatible, which requires that 1a tell the truth, we must have $f(1a, 2a) = \alpha$. But then type $2a$ would strictly like to pretend to be type $2b$: by telling the truth, he would get α for sure, so that his expected utility from truth-telling is

$$\text{prob.}(1a, 2a) \cdot 0 + \text{prob.}(1b, 2a) \cdot 2 = 1,$$

whereas his expected utility from lying, which ensures that β is implemented, would be

$$\text{prob.}(1a, 2a) \cdot 6 + \text{prob.}(1b, 2a) \cdot 0 = 3.$$

In this example, no classically efficient decision rule can be incentive compatible!

Holmström and Myerson argue that only three (one for each evaluation stage) of the above six efficiency notions are relevant: ex ante incentive efficiency Δ_A^*, interim incentive efficiency Δ_I^*, and ex post classical efficiency Δ_P. Indeed, if the entire information state θ were to become publicly known before the decision in A is chosen, then there would be no incentive problems and Δ_P would be the right efficiency concept to use. If the decision rule must be selected when each individual knows only his own type, then the incentive compatibility must be required, and Δ_I^* is the right efficiency concept. If the decision rule can be selected before the individuals learn their types, but if the individuals cannot commit themselves ex ante to honestly report their types after they learn them, then Δ_A^* is the appropriate efficiency concept.

Holmström and Myerson then point out that, if a decision rule f is incentive efficient (in the interim sense) then a social planner who does not know any individual's actual type could not propose any other incentive-compatible decision rule that every individual in any type is sure to prefer. However, there could possibly exist another incentive-compatible rule g and an information state θ such that

$$(\mathrm{G}) \quad U_i(g|\theta_i) > U_i(f|\theta_i), \ \forall i.$$

In this case, if θ were the actual information state, then all individuals would unanimously prefer g over f, each given his respective type.

Such unanimity may not be effective for replacing f, however. The problem is that, even if (G) holds in state θ, it may be that individual 1 would reverse his preference to favor f if he learned that individual 2 also preferred g over f, since agent 2's preference would reveal new information to agent 1 about agent 2's type. If the agents were to unanimously agree to change from f to g, then it would be common knowledge (in the sense of Aumann) that all individuals prefer g over f.

Thus Holmström and Myerson continue to ask whether it could be common knowledge that all the individuals in the economy prefer g over f, when each individual knows only his own type. To this end, they define common-knowledge events such that θ is contained in the common-knowledge event R, then all I agents assign probability zero to the states outside of R. Holmström and Myerson say that g interim dominates f within a common-knowledge event $R \neq \emptyset$ if and only if (G) holds at each and every $\theta \in R$ with at least one strict inequality. They then reach a conclusion:

Theorem HM-1. An incentive-compatible decision rule f is interim incentive efficient if and only if there does not exist any common-knowledge event R such that f is interim dominated within R by another incentive-compatible decision rule g.

Theorem HM-1 implies that, if f is incentive efficient and each individual knows only his own type, then it cannot be common knowledge that the individuals unanimously prefer some other incentive-compatible decision rule g over f. Indeed, if such a decision rule g exists then one can define a decision rule h, which coincides with g in event R and with f outside of R, and one can verify that h is incentive feasible, and it dominates f in Δ_I^*, which is a contradiction.

Theorem HM-1 does not mean that the individuals could never reach a unanimous agreement to replace f by some other incentive-compatible decision rule. It only means that if a unanimous agreement is reached then each individual must know more than just his own type; communication must have occurred. With the latter type of information leakage, it may happen that no ex-ante incentive efficient game form can remain durable.

Let us look at an example. Suppose that there are two individuals in the economy, and each individual may be one of two possible types. Individual 1 may be type la or lb, individual 2 may be type 2a or 2b, and all four possible combinations of types are equally likely. There are three possible decisions called α, β, and γ. The utility payoff of each individual from each decision depends only on his own type, as shown in the following table.

	u_{1a}	u_{1b}	u_{2a}	u_{2b}
α	2	0	2	2
β	1	4	1	1
γ	0	9	0	-8

In this example, individual 2 in either type and individual 1 in type 1a both prefer α over β and β over γ. However if individual 1 is type 1b then his preference ordering is reversed and he strongly prefers γ. Type 2b differs from 2a in that 2b has a greater aversion to decision γ. (These payoffs are von Neumann-Morgenstern utility numbers.)

Among all incentive-compatible decision rules, the following decision rule f uniquely maximizes the sum of the two individuals' ex-ante expected utilities:

$$f(1a, 2a) = \alpha,\ \ f(1b, 2a) = \gamma,$$

$$f(1a, 2b) = \beta,\ \ f(1b, 2b) = \beta.$$

Notice that this decision rule selects decision γ, type 1b's most preferred decision, if the types are 1b and 2a; but if 2's type is 2b (so that 2 is more strongly averse to γ) then the decision rule selects β instead.

To check that f is incentive compatible, notice that type 2a can get decisions α or γ with equal probability if he is honest, or he can get β for sure if he lies and reports his type as 2b. Since both of these prospects give the same expected utility to 2a, he is willing to report his type honestly when f is implemented. This decision rule f is incentive efficient (in both the interim and ex ante senses), so no outsider could suggest any other incentive-compatible decision rule that makes some types better off without making any other types worse off than in f.

But if individual 1 knows that his type actually is 1a, then he knows that he and individual 2 both prefer decision α over this decision rule f. Thus, rather than let f be implemented, individual 1 in type 1a would suggest that decision α be implemented instead, and individual 2 would accept this suggestion. Thus, although f is an incentive-efficient decision rule, it is possible for the individuals to unanimously approve a change to some other decision rule (namely α-for-sure).

Of course, this unanimity in favor of α over f depends on 1's type being 1a, but consider what would happen if 1 were to insist on using f rather than α. Individual 2 would infer that l's type must be 1b. Then decision rule f would no longer be incentive compatible, because both types of individual 2 would report "2b", to get decision β rather than γ.

Thus, if the individuals can redesign their decision rule when they already know their own types, then the decision rule f could not be implemented in this example, even though it is incentive efficient. Holmström and Myerson thus call f incentive efficient but not durable.

Despite the above example, Holmström and Myerson present a class of durable incentive efficient decision rules. To specialize, they assume that no information state in Θ has zero probability, and they say that an incentive-compatible decision rule f is *uniformly incentive compatible* if and only if, when expecting the other agents $j \neq i$ to always report their types honestly, no agent i would ever want to report $\hat{\theta}_i \neq \theta_i$, even if when reporting $\hat{\theta}_i$, agent i has learned about the realization of

θ_{-i}. For example, a decision rule that selects a constant decision in A independently of θ would be uniformly incentive compatible.[36]

Theorem HM-2. If f is uniformly incentive compatible and interim incentive efficient, then f is durable.

A special case is this: suppose that individual 1 is the only individual with any private information, so that every other individual has only one possible type. Then there are no incentive constraints for the individuals other than agent 1; and agent 1 already knows all the others' types when he reports his type into a decision rule. Thus *every* incentive-compatible decision rule is uniformly incentive compatible in this case, and the following result obtains.

Theorem HM-3. If there is only one individual with private information then every interim incentive-efficient decision rule is durable.

[36]If every individual's utility function is independent of the other agents' types, then uniform incentive compatibility is equivalent to honesty being a dominant strategy for every individual.