The not-so-secret-agent:
professional monitors, hierarchies and implementation.

Sandeep Baliga*
Forthcoming in *Review of Economic Design*

* Kellogg GSM (M.E.D.S.), 2001 Sheridan Road, Evanston IL 60208-2009.
This is Chapter 2 from my Ph.D. at Harvard University. Many thanks to my advisors Jerry Green, Eric Maskin and Tomas Sjöström for their help. Serge Moresi, Anna Paulson, Michael Riordan, an Assistant Editor and two referees also gave very useful comments. Any remaining errors are mine.

Abstract

It is well-known that, when agents in an organization possess private information that is unverifiable by an outside party, games where agents simply announce their information can have multiple equilibria that may impede the successful implementation of the organization's objectives. We show that the introduction of a professional monitor (e.g. auditor, regulator, supervisor) can help to destroy the "bad" equilibria when agents have private information but have incomplete information about others' information.

1. Introduction

A central authority or principal such as an entrepreneur or government deals with agents who have private information that is relevant for optimal decisions. For example, a planner is concerned with agents' preferences over production of some public good in order to decide how much to provide. The costs of different electricity generation plants are pertinent when the government determines how to allocate funds and output among them. The optimal investment strategy of shareholders in a multi-product firm depends on the profitability of different product-based divisions.

The central authority can set up a procedure to collect the information agents possess. In a model of public good production, d'Aspremont and Gerard-Varet (1979) construct such a game form or mechanism where agents announce their preferences (a "direct revelation game"), truthtelling is necessarily a Bayesian-Nash equilibrium (i.e. each agent weakly prefers to tell the truth whatever his information given that all other agents always tell the truth), and this equilibrium implements the first-best level allocation of goods. However, it is possible that direct revelation games can have other equilibria where agents lie. Such equilibria may not be optimal from the perspective of the central authority [1].

In these situations, we often observe an agent who has been employed by some central authority to monitor other agents. The Inland Revenue Service which audits tax returns, industry regulators and outside appointees on the boards of directors of firms are examples of this phenomenon. These agents obtain information to help alleviate the asymmetric information problem

[1] See Example 23.AA.1 in Mas-Colell, Whinston and Green (1995) pages 911-912. They show how the multiple equilibrium problem can occur in this setting. See Palfrey and Srivastava

(see for example Laffont and Tirole (1991) on the use of information collected by regulators to monitor firms). However, we show that, even if these professional monitors have no such information collection role, their introduction into organizations can in part be understood from the perspective of *full* implementation of the objectives of the central authority (the social choice function (scf)). In other words, their policing role in mechanisms guarantees that *all* equilibria are optimal ("full" implementation) rather than just some ("partial" implementation). We focus on the full implementation with incomplete information case where the professional monitor is policing agents who have private information that is unobservable to other agents.

For this interpretation of professional monitors to be consistent, we must assume that the scf does not change when the extra agent is introduced into the organization. Otherwise, we would simply add another agent into the implementation problem. In other words, the professional monitor is an instrument used by the central authority to fully implement its objectives and does not directly affect its objectives. Therefore, we are allowing the central authority not only the ability to design message spaces and an outcome function, as in standard implementation theory, but also to add a professional monitor. This strategic possibility is actually exploited in many applied settings: The Inland Revenue Service is an instrument used by the government to implement its objectives and does not affect them directly. The government's social objectives regarding resource and output allocation among different electricity generating plants does not directly depend on the presence or absence of a regulator. Shareholders appoint outsiders onto the board of directors and their optimal investment strategy is independent of presence of these observers.

Notice that the notions of full and partial implementation are defined with respect to some

---

(1987) for other examples.

concept of equilibrium. It is well known that equilibrium analysis requires strong assumptions on players' rationality and beliefs. Therefore, we see our analysis applying when the professional monitor has a small number of agents he monitors on a long-term basis. In our examples above, the tax auditor should have a small number of taxpayers he regularly deals with, the regulator a few electricity generating plants he oversees for a significant period and the outside director should be on the board long enough to understand the firm's position.

Our mechanism is simpler than standard mechanisms in the implementation literature as it does not require the use of the so-called "integer-games."[2] For example, Palfrey (1992), while he was the first to introduce the idea that an uniformed agent can facilitate full implementation, employs such a game as do many other authors (see Palfrey and Srivastava (1989)). Such constructions are interesting as existence theorems to characterize the set of implementable allocation rules but are not practical, "real-world" prescriptions for mechanism design.

Moreover, as the mechanisms analyzed in the Bayesian implementation literature have only one stage, they can only implement optimal plans that satisfy the necessary conditions for Bayesian implementation including the Bayesian monotonicity condition[3] and the selective elimination condition (Mookherjee and Reichelstein (1990)). Roughly speaking, these conditions require that (a) when agents play non-optimally, there should exist some agent of some type who strictly prefer some (possibly random) outcome from a deviation to the outcomes identified by the social choice function and (b) that this agent of this type should prefer the outcomes identified by the social choice

[2] In mechanisms employing integer games, agents' strategy sets include integers and the outcome function specifies that the agent who announces the highest integer is allowed to choose his favorite outcome in certain circumstances.

[3] As we would have to introduce a lot of notation to define this condition, we simply refer the

function to the outcomes following the deviation when agents play optimally. That is, there should be "preference reversal" around the scf. By using a dynamic mechanism and a not-so-secret agent, we allow preference reversal with respect to outcomes not identified by the social choice function[4]. Therefore, we dispense entirely with the Bayesian monotonicity and the selective elimination conditions.

Moore and Repullo (1988) present a case where they can avoid the use of integer games when there is complete information. However, our assumption of incomplete information is the more interesting and relevant one if we are interested in decentralizing information as much as possible.

Let us give the intuition for our result using a toy example before going on to a literature survey. (For later reference, the reader should bear in mind that the details of the example will be used to give intuition for the notation we use in the next section.) Suppose there are two agents, Bob and Charles, who can either be of low or high ability and their abilities are independently and identically distributed. The entrepreneur is interested in their abilities as she wants to design a production line that depends on this information. Workers know their own ability but only know the prior distribution over the other's abilities. Abilities differ only in that there is so-called "preference reversal" over two outcomes: a high ability worker strictly prefers doing overtime for $100 rather than a day off while a low ability worker strictly prefers the reverse. These outcomes are used in the mechanism described below to "prove" if a worker is high ability or low ability.

There is a manager, Andrew, employed by the entrepreneur to monitor the workers. Andrew

---

interested reader to Palfrey's (1992) excellent survey.

[4]A companion paper, Baliga (1997), provides a sufficient condition for sequential implementation in

does not know either Bob's or Charles's abilities but, as in standard principal-agent theory and equilibrium analysis, he does know the chosen strategy profile. The entrepreneur designs the following two-stage game of observed actions. In such games, all players choose actions simultaneously in each stage and the action profile is revealed at the end of each stage. In Stage 1, the two workers and Andrew simultaneously announce "messages" which are revealed at the end of the stage. In particular, workers announce their ability and Andrew either agrees or challenges by picking on a worker (his victim) and announcing an ability for the worker. If Andrew agrees or Andrew's announcement of a worker's type is the same as the worker's own one, the entrepreneur designs the production line as if the workers' announcements are true; if Andrew challenges and his announcement for his victim's type differs from that of his victim's own announcement, we fine his victim and go to Stage 2.

In Stage 2, Andrew's victim, say Bob, chooses between overtime or a day off. Notice that Bob prefers overtime to a day off if and only if he is high ability. Therefore, if he announces he is low ability (respectively, high ability) in Stage 1 and chooses overtime (respectively, a day off) in Stage 2, we say his Stage 2 choice contradicts his Stage 1 announcement. If Bob's choice contradicts his Stage 1 announcement, Andrew is rewarded by giving him the fine we imposed on Bob. If there is no contradiction, we fine Andrew. This system of rewards and fines gives Andrew the incentive to challenge whenever the workers are lying. In Stage 1, workers will then try to avoid being fined by announcing the same ability as Andrew. In fact, the only strategy profile where there is no incentive for Andrew to challenge or for a worker not to match Andrew's announcement of his type is the one where Bob and Charles both tell the truth because, if Andrew challenges in this case,

---

an abstract setting.

he will sometimes get fined in Stage 2. Incentive compatibility has to be imposed on the entrepreneur's optimal plan to guarantee that worker's will in fact tell the truth in equilibrium. Therefore, the employment of a professional monitor helps to get rid of the "bad" equilibria and retain the "good" ones.

In the next Section, we present our main result. Sections 3 concludes.

2. Adverse selection

Suppose there are two agents. (The mechanism we present can easily be generalized to the case of more than two agents.) Their preferences are indexed by their type θ_i ($i = 1,2$) which is drawn from a set Θ_i ($i = 1,2$). We assume these sets are finite. Let $d \in D$ be a central decision and $t=(t_1,t_2) \in \Re^2$ a profile of transfers of a private good from the two agents. Let T denote the set of *feasible* transfers in that $t_1+t_2 \leq 0$. The agents' von Neumann-Morgenstern utility functions are

$$v_i : D \times \Re \times \Theta_i \rightarrow \Re \text{ where } v_i((d,t_i),\theta_i) = u_i(d,\theta_i) - t_i \; (i = 1,2)$$

and $u_i : D \times \Theta_i \rightarrow \Re$. We remark that $d \in D$ can be interpreted as an allocation of private and/or public goods.

Each agent knows his own type but not that of the other agent. Note that we are making a "private values" assumption: an agent's preferences depend on his own type and not those of the other player[5]. The prior probability of player i of type θ_i is $p_i(\theta_i)$. These probabilities are positive for all θ_i in Θ_i and for all i and the prior distribution is common knowledge. Conditional beliefs $p_i(\theta_{-i} | \theta_i)$

5 An alternative assumption is "common values": each agent's preferences are allowed to depend on the type of the other agent. We do not consider this case in this paper.

are derived from the prior using Bayes' rule. We assume that two types of a given player i, θ_i and φ_i say, differ in that there are at least two central decision and transfer pairs, (x,t_x) and (y,t_y) over which their preferences are reversed in the following sense:

Assumption (PR) $\quad u_i(x,\theta_i)-t_x > u_i(y,\theta_i)-t_y$ but $u_i(y,\varphi_i)-t_y > u_i(x,\varphi_i)-t_x$.

In our toy example, this *preference reversal* condition is satisfied with the two types as high and low ability and two outcomes (overtime,$100) and (day off, $0).

Let us turn to the not-so-secret agent, 007. We assume that agent 007 has a von Neumann-Morgenstern utility, $u_{007}(t_3)$ ($u_{007} : T_3 \rightarrow \Re$, $T_3 = \Re$) that does not depend on the central decision, is strictly increasing in the transfer t_3 to him, and is unbounded above. (We argue later that the first part of this assumption can also be relaxed.) We normalize 007's reservation utility to zero for simplicity.

The not-so-secret-agent is part of the means used by a central authority to achieve its ends. Therefore, he does not enter our notion of social welfare: A social choice function, $f: \Theta_1 \times \Theta_2 \rightarrow D \times T$, gives the central decision and feasible transfers for each preference profile (θ_1,θ_2). The central decision is $d = D(\theta_1,\theta_2)$ and transfers $(t_1,t_2) = (T_1(\theta_1,\theta_2),T_2(\theta_1,\theta_2))$ for a profile (θ_1,θ_2). In the toy example, the central decision was the optimal production line, the transfers were the wages and the preference profile was the ability profile of Bob and Charles. We require that the social choice function be Incentive Compatible.

Definition The social choice function, $f = (D,T_1,T_2)$, satisfies *Incentive Compatibility* (*IC*) if and only if for all $i \in \{1,2\}$ and for all θ_i in Θ_i , we have

$$\sum_{\theta_{-1}\in\Theta_{-1}}[u_i(D(\theta_i,\theta_j),\theta_i)-T_i(\theta_i,\theta_j)]p_i(\theta_{-i}\mid\theta_i)\geq\sum_{\theta_{-1}\in\Theta_{-1}}[u_i(D(\theta'_i,\theta_j),\theta_i)-T_i(\theta'_i,\theta_j)]p_i(\theta_{-i}\mid\theta_i)$$

for all $\theta'_i \in \Theta_i$.

This is a well-known necessary condition for implementation in incomplete information environments.

We now define some variables that will be useful in the statement of our theorem. Let $p^* = min_{i,\theta_i}p_i(\theta_i)$, $i = 1,2$. For a given $\varepsilon>0$, choose a real number t^{007} such that

(*) $p^*u_{007}(t^{007})-(1-p^*)\varepsilon > 0$.

In the mechanism we describe below, the transfer t^{007} will be used to reward agent 007 if he catches an agent lying and we will fine him ε if he claims that an agent is lying but it turns out that he is not. Expression (*) guarantees that even if an agent is lying with the smallest possible probability p^*, the reward for catching him is large enough to compensate in expectation the professional monitor for getting fined with probability $(1-p^*)$.

We define some notation before stating our main result: For $v\in\{1,2\}$, let $G_v:\Theta_v^2\rightarrow(D\times\Re)\times(D\times\Re)$ be such that for all θ_v and φ_v, if $\theta_v\neq\varphi_v$, $((x,t_x),(y,t_y)) = G_v(\theta_v,\varphi_v)$ satisfies:

$u_v(x,\theta_v)-t_x > u_v(y,\theta_v)-t_y$, and

$u_v(y,\varphi_v)-t_y > u_v(x,\varphi_v)-t_x$.

Our assumption (PR) ensures that such functions can be found. We will write $G_v = G_{xv}\times G_{yv}$ (v={1,2}). The boundedness of utility functions is as follows. For $i = 1,2$ let R_i denote the range of (D,T_i) from the domain $\Theta_1\times\Theta_2$ and for $v = 1,2$ let R^*_{xv} denote the range of G_{xv} from the domain Θ_v^2

and R^*_{yv} denote the range of G_{yv} from the domain Θ_v^2 . Then for $i = 1,2$, as there are a finite number of types by assumption,

(**) $u_i(d,\theta_i) - t_i$ is uniformly bounded for all θ_i in Θ_i for all (d,t_i) in $R_1 \cup R_2 \cup R^*_{x1} \cup R^*_{y1} \cup R^*_{x2} \cup R^*_{y2}$.

While expression (*) concerns incentives for the professional monitor, expression (**) concerns incentives for the agents. It guarantees that a fine, Δt, can be found such that any agent forced to pay the fine strictly prefers to have some (d,t_i) in $R_1 \cup R_2 \cup R^*_{x1} \cup R^*_{y1} \cup R^*_{x2} \cup R^*_{y2}$ implemented rather than pay the fine.

We now describe the *professional monitor, Δt-mechanism*, we use to prove the theorem. It is a multi-stage game of observed actions. That is, in each stage all three players simultaneously choose an action that is revealed at the end of the period. Also, the set of feasible actions can be dependent on the stage and on history (see Fudenberg and Tirole (1991), page 331, for a detailed description of this class of games).

<u>Stage 1</u>: All players simultaneously announce their own types (call the ensuing vector of types θ) and agent 007 simultaneously either agrees or challenges by picking a player v ($v = 1,2$) and a type φ_v.

There are three cases to consider:

(1) if agent 007 agrees, implement $[d(\theta),t_1(\theta),t_2(\theta)]$;

(2) if agent 007 challenges and $\theta_v = \varphi_v$, implement $[d(\theta),t_1(\theta),t_2(\theta)]$; and

(3) otherwise go to Stage 2.

The message profile is revealed at the end of the stage.

Stage 2: Player v picks between $(d_{1v},t_{1v}) = (x,t_x+\Delta t)$ and $(d_{2v},t_{2v})=(y,t_y+\Delta t)$ where $((x,t_x),(y,t_y)) = G_v(\theta_v,\varphi_v)$. If player v chooses $(x,t_x+\Delta t)$, player $w \neq v$ gets $\Delta t+t_x+F$ where F is the fine the agent 007 has to pay to reduce his utility by ε as punishment. If he chooses $(y,t_y+\Delta t)$, agent 007 gets t^{007} and player w gets $\Delta t+t_y-t^{007}{}_v$.

The *set of feasible outcomes* of the professional monitor, Δt-mechanism is $X \equiv (D,T_1,T_2,T_3)$ with generic element $x=(d,t_1,t_2,t_3)$. Notice that all the final outcomes identified by the mechanism satisfy budget balance - no money is ever thrown away. This desirable property of a mechanism is also a concern of d'Aspremont and Gerard-Varet (1979) and is not satisfied by Groves mechanisms (see Chapter 7 of Fudenberg and Tirole (1991) for an analysis of such mechanisms and for an extensive list of references).

The solution concept we use is perfect Bayesian equilibrium: A *perfect Bayesian equilibrium* is a strategy profile and beliefs pair, (σ,μ), where (1) (sequential rationality) strategies are optimal for all agents at all their information sets given their beliefs and (2) beliefs are updated using Bayes' rule wherever possible.

Let PB be the set of perfect Bayesian equilibrium outcomes of the professional monitor, Δt mechanism and let $PB(\theta)$ be the set of such outcomes for preference profile θ. The mechanism *implements* a scf $f=(D,T_1,T_2)$ if $PB(\theta) = (D_1(\theta),T_2(\theta),T_3(\theta),0)$ for all $\theta\in\Theta$. We are now ready to state our main result.

Theorem. Suppose (*) and (**) are satisfied. For any incentive compatible scf, there exists a Δt

such that the scf can be implemented in perfect Bayesian equilibrium by the professional monitor, Δt-mechanism.

Let us describe the mechanism verbally and analyze it intuitively before the formal proof. In Stage 1, players announce private information, θ_1 and θ_2, and 007 either accepts or challenges by picking a player, v, and a type for the player, φ_v (all these announcements are made simultaneously). We go to Stage 2 if and only if 007 challenges and announces a type φ_v that differs from player v's own announcement, θ_v. Otherwise, the mechanism acts as if the two players told the truth and implements an outcome that is optimal given their messages. If we enter Stage 2, 007's victim, player v, pays a fine Δt and is then forced to choose between two outcomes (x,t_x) and (y,t_y) where player v strictly prefers (x,t_x) over (y,t_y) if he is indeed of type θ_v but the reverse if he is of type φ_v. If by his choice of (y,t_y) player v reveals he is not of type θ_v, 007 is rewarded by paying him t^{007}; otherwise, 007 is fined F. To ensure balancedness, any remaining transfer from player v is transferred to player $w \neq v$.

Suppose players 1 and 2 play strategies such that at least one type for some player, say type φ_1 of player 1, does not tell the truth. If the mechanism ever goes to Stage 2, player 1 of type φ_1 picks the appropriate outcome. If this is an equilibrium, agent 007 knows that if he challenges player 1 and announces φ_1, he will get rewarded with positive probability. Expression (*) implies that this is a strategy he strictly prefers to agreeing. Is a strategy profile where player 1 and 2 do not always tell the truth and 007 challenges an equilibrium? Notice that in such a strategy profile some agent of some type must be paying Δt. If in Stage 1 he deviates and sends the message that equals 007's challenge, he can prevent the mechanism going to Stage 2 and avoid the fine. Expression (**)

guarantees that if Δt is large enough, this he strictly prefers to deviate in this way so the original strategy profile is not an equilibrium. Only truthtelling on the part of player 1 and 2 and agreeing on part of 007 is an equilibrium: players 1 and 2 have no incentive to deviate as the scf is incentive compatible.

Proof. See the Appendix.

3. Extensions and Conclusions

This paper has shown that adding a professional monitor to an organization when information is decentralized ensures that the information can be collected so the organization's objectives are always met. This professional monitor may do no productive work himself but guarantees that there are no non-optimal equilibria.

We can allow more general specifications of agent 007's utility function, including uncertainty over his preferences, as long as it is bounded over the same domain as those of players 1 and 2 and he continues to strictly prefer more money to less. We can then define rewards that are large enough so that any type of agent 007 strictly prefers to sneak on a potential liar rather that agree to a deception on the part of the agents. What is not possible is to include agent 007's information in the scf because then we obtain a three agent model rather than a two agent one. While another agent can be added to monitor these three, this may begin to stretch our assumption that our analysis applies only to small groups. This in turn suggests that we may have the germs of a theory of hierarchies of the following form: At the lowest level are workers who do only productive work. Supervisors are employed to monitor them in the sense above (i.e. to ensure full implementation) but may in turn also do productive work. Managers are one level above supervisors and monitor them and also do productive work. We can continue in this way adding an extra tier at each stage till we reach a final tier consisting of the owners who offer no direct input into the production process. An assumption that common knowledge conditions sufficient to justify equilibrium analysis apply only to small groups is needed to derive such a pyramidical structure: otherwise we can just add one extra super-monitor to supervise all productive agents. Therefore, the development of this theory and the extension of our model depends on an analysis of the interaction of the size of groups where equilibrium analysis applies with implementation considerations. This is an interesting line of research to pursue in the future.

4. Appendix

The Appendix contains the proof of Theorem 1.

Proof of Theorem 1. We prove two claims: first, there does not exist a pure strategy perfect Bayesian equilibrium where either player 1 or 2 or both play any strategy other than truthtelling in Stage 1 so there are no non-optimal equilibria; second, it is indeed a perfect Bayesian equilibrium of our mechanism for players 1 and 2 to tell the truth and agent 007 to agree and this equilibrium implements the scf. The two claims together prove the theorem.

Claim 1. Any strategy profile where either player 1 or 2 do not tell the truth is not part of a perfect

Bayesian equilibrium of the mechanism above.

Proof. Suppose not and consider any strategy profile where player 1, say, sends Stage 1 messages where he does not always reveal his type truthfully. Then there exists at least one type of player 1, say θ_1, who announces θ'_1 and strictly prefers the appropriate $(y,t_y+\Delta t)$ corresponding to θ'_1 and θ_1 to the appropriate $(x,t_x+\Delta t)$. There may also be a type θ''_1 of player 1 announcing $\delta_1\neq\theta''_1$, who strictly prefers the $(y,t_y+\Delta t)$ corresponding to δ_1 and some $\theta'_1\neq\theta''_1$ even though player 1 of type θ'_1 is announcing truthfully.

Notice from our definition in assumption PR that player 1's preferences over $(x,t_x+\Delta t)$ and $(y,t+\Delta t)$ do not depend on his beliefs over player 2's type. Hence, even if agent 007's challenge sends the game "off-the-equilibrium-path", the beliefs we allocate for player 1 in Stage 2 do not matter. This in turn implies that, given the equilibrium strategy, if in Stage 1 agent 007 puts positive probability on player 1 lying, he puts positive probability on agent 1 picking the relevant $(y,t_y+\Delta t)$ in Stage 2. By (*), agent 007, therefore, gets positive expected utility from challenging in this way while he gets zero from agreeing or challenging and announcing a type of player 1 that is never contradicted by his Stage 1 messages (the latter situation could arise if all types of player 1 are announcing the same type, θ'_1, in Stage 1 and agent 007 challenges by announcing θ'_1). Therefore, he will challenge in any equilibrium where agent 1 does not always tell the truth.

Let us turn to the strategy of the types of player 1 not announcing φ_1. If they announce φ_1, some outcome in the range of the scf will be implemented as the second case in Stage 1 will apply. Given (**), we can choose Δt so large that all types of player 1 prefer the outcome of the scf if they announce φ_1 or the outcome if player 2 is picked by agent 007 to not announcing truthfully in Stage

1. This contradicts our supposition that player 1 was playing an equilibrium strategy. Hence, he must be truthtelling in any perfect Bayesian equilibrium. A similar argument to the one above shows that player 2 will then also play optimally with respect to the scf. Claim 1 is proved.

Claim 1 shows there are no non-optimal equilibria.

Claim 2. A perfect Bayesian equilibrium[6] of the mechanism above is for players 1 and 2 to tell the truth in Stage 1 whatever their type and to pick the appropriate $(x,t_x+\Delta t)$ in Stage 2, and for agent 007 to agree (any beliefs can be allocated in Stage 2 which is off the equilibrium path). This equilibrium implements the scf.[7]

Proof of Claim 2. Given that players 1 and 2 are telling the truth, agent 007 strictly prefers to agree as he places zero probability on the event that he will get rewarded for catching a liar and he loses ε with positive probability for accusing someone wrongly. Given that IC holds and agent 007 is agreeing, truthtelling is an equilibrium strategy in Stage 1 for players 1 and 2. Given that they have

---

6 This is in fact the unique pure strategy perfect Bayesian equilibrium of our mechanism. In general, full implementation of social choice functions does not require uniqueness of equilibrium strategies but uniqueness of equilibrium outcomes.

7 An alternative solution concept is sequential equilibrium (Kreps and Wilson (1982)). Any sequential equilibrium is also a perfect Bayesian equilibrium. Hence, in Claim 1 when we show that some strategy-beliefs pairs are not perfect Bayesian equilibria, we also show that they are not sequential equilibria. It only remains to show that the equilibrium identified in Claim 2 satisfies the consistency condition (for brevity we refer the reader to Kreps and Wilson (1982) for a definition). Choose a sequence of totally mixed strategies as follows: let all players put probability $(1-\lambda)$ on the equilibrium strategies and a total probability of λ on all other strategies at each point in the game. As we allow λ to converge to zero, the Bayes' consistent belief of agent 007 is to assume the preference profile announced in Stage 1 is true.

told the truth in Stage 1, players will choose the relevant $(x,t_x+\Delta t)$ in Stage 2 as this is optimal given their preferences. As the scf satisfies incentive compatibility this equilibrium implements the scf. Claim 2 is proved. *Q.E.D.*

5. References

D'ASPREMONT, C. AND GERARD-VARET L. A. "Incentives and Incomplete Information." *Journal of Public Economics*, Vol. 11 (1979), pp. 25-45.

BALIGA, S. "Implementation in Economic Environments with Incomplete Information: the Use of Multi-Stage Games," forthcoming , *Games and Economic Behavior*.

FUDENBERG, D. AND J. TIROLE . *Game Theory.* Cambridge: MIT Press, 1991.

KREPS, D. AND WILSON R. "Sequential Equilibrium." *Econometrica*, Vol. 50 (1982), pp. 863-894.

LAFFONT, J.-J. AND TIROLE J. "The Politics of Government Decision-Making: A Theory of Regulatory Capture, " *Quarterly Journal of Economics*, Vol. 106 (1991), pp. 1089-1127.

MAS-COLELL, A., WHINSTON M. AND GREEN J. *Microeconomic Theory*. Oxford: Oxford University Press, 1995.

MOOKHERJEE, D. AND REICHELSTEIN S. "Implementation via Augmented Revelation Mechanisms." *Review of Economic Studies*, Vol. 57 (1990), pp. 453-475.

MOORE, J. AND R. REPULLO R. "Subgame Perfect Implementation," *Econometrica*, Vol. 56 (1988), pp. 1191-1220.

PALFREY, T. "Bayesian Implementation", in *Advances in Economic Theory: Sixth World Congress of the Econometric Society*, Volume I (J.J. Laffont, ed.), Cambridge: Cambridge University Press, 1992.

PALFREY, T. AND SRIVASTAVA S. "Implementation with Incomplete Information in Exchange Economies." *Econometrica*, Vol. 57 (1989), pp. 115-134.