Absence of Commitment in Principal-Agent Games

R. Vijay Krishna*
Department of Economics
Pennsylvania State University
University Park, PA 16802

Extended Abstract

The Revelation Principle is the cornerstone of the theory of mechanism design. It shows that seemingly complex problems can be viewed as programming problems with constraints that give the agent the right incentives to truthfully reveal his private information. Needless to say, the theory is widely applicable. Nevertheless, one of the main assumptions underlying the Revelation Principle is not satisfied in many applications. This is the assumption that the Principal can commit to any outcome of the mechanism. If this assumption is not satisfied, the set of outcomes is (usually) significantly smaller.

A major advance in the study of Principal-Agent games without commitment is due to Bester and Strausz [3]. They show that even with imperfect commitment, the problem can be reduced to a programming problem, albeit under slightly stronger conditions. Nevertheless, the set of outcomes is (usually) significantly smaller than what could have been obtained if the Principal had been able to commit, or equivalently, there had been a mediator. It should be noted that another advancement due to Bester and Strausz is in formulating the problem so that all cases between perfect commitment and no commitment can be handled simultaneously. However, they rely on only one round of communication. Recently, Aumann and Hart [1] have shown that with repeated communication, the set of equilibrium outcomes can once again be extended and provide a complete characterisation (in terms of bimartingales and biconvexity) of the set of outcomes thus obtained. Once again, the set of outcomes is smaller than the set of outcomes with perfect commitment or a mediator.

In this paper, we shall show that we can virtually attain strictly more outcomes even in the absence of a mediator. We draw on the similarities alluded to in Myerson [4], where he shows that noisy communication mechanisms can help agents replace a mediator in games of perfect information (i.e. achieve the set of correlated equilibria) and in Sender-Receiver games. We adapt a mechanism due to Ben-Porath [2] to achieve this. The key to a noisy mechanism, in the case of Sender-Receiver games for example, is the observation that the receiver of a message should not be sure just which type of agent has sent him the message, even after accounting for the fact that the various Sender types may be randomising and using Baye's rule to update probabilities. This is precisely the service that a mediator provides. We make this precise below.

*Contact Information: 608, Kern Graduate Building, University Park, PA 16802. Tel: + (814) 237 1280, Email:vijayk@psu.edu

Let A be the set of payoff-irrelevant actions that the Sender can take. Let X be the set of signals that the Receiver gets. Let $(X, \mathcal{X})$ be a measurable space and (with a slight abuse of notation), $\Delta(X)$ denote the space of probability measures on this measurable space.

Definition 1. *A* **communication mechanism** *is a function* $\mu : A \to \Delta(X)$.

Thus, the signal that the Receiver gets is a random variable drawn according to the probability measure $\mu(a)$ over $(X, \mathcal{X})$.

Condition T. A communication mechanism satisfies Condition T iff for all $a \in A$, $\text{supp}(\mu(a))$ is a singleton.

Definition 2. *A communication mechanism is* **noisy** *if it does not satisfy Condition T.*

Clearly, a mediated mechanism is an example of a noisy communication mechanism. Another example of a noisy communication mechanism is in Myerson [4] where the message is carried by a pigeon that could get lost with some probability. Classical cheap talk is an example a noiseless communication mechanism. The mechanism we provide is noiseless in precisely such a sense. However, the complete mechanism is a pasting together of this mechanism with a another protocol which depends on jointly controlled lotteries.

A mediator provides a similar function in games of perfect information, where the players want to implement the set of correlated equilibria. Ben-Porath shows that (in finite games) if there are two distinct Nash equilibrium payoffs for each player then any rational correlated equilibrium with payoffs above these two payoffs can be virtually implemented in sequential equilibrium. Ben-Porath's result holds for any number of players. However, the mechanism for two players is a little more involved. It involves players exchanging messages through urns where the contents of the urn can be publicly verified and verification actually takes place with some probability which is modelled as the outcome of a jointly controlled lottery.

We adapt Ben-Porath's techniques in this paper by using a *menu* of urns. We first consider the case of games with atleast two equilibria in the Sender-Receiver game where the payoffs to the Receiver are distinct. Call these payoffs v_1 and v_2. Consider an equilibrium of the mediated game with payoffs to the Receiver above v_1 and v_2. Suppose there is a finite set of types T and for all types $t \in T$, the mediator induces tells the Receiver what action to take with probability $\mu(\cdot|t)$. As this is an equilibrium, this distribution satisfies the incentive compatibility conditions. Now in order to replace the mediator, we let the Receiver choose a menu of $|T|$ urns, with the contents of each urn being a set of identical balls, whose contents, in turn, are an action to be taken by the Receiver and cannot be seen by the Sender. The Receiver provider the Sender with a menu which coincides with the the distribution prescribed by $\mu(\cdot|t)$ and labels the urns according to the types of the Sender. The Sender picks a ball from one of the urns and gives it to the Receiver who opens the ball (which only he can do) and performs the prescribed action. Prior to performing the prescribed action, both players destroy all the other urns and balls. It is easy to see that it is incentive compatible for a Sender of type t to pick a ball from the urn labelled t. This is because the distribution of actions as per the balls is supposed to mimic the distribution that a mediator would induce. The only thing to check is the Receiver's incentives to provide the correct distribution of balls in the urns. In order to verify this, we let the players this perform this process repeatedly and with some probability, the contents of all the urns are checked. This probability is the outcome of a jointly controlled lottery. If the Receiver has deviated, then they play the worst outcome which is the Babbling equilibrium. If the Receiver has not deviated, then they go through the entire process again. This happens finitely many times and in the last period, if there is no deviation, they play some distribution over v_1 and v_2. We thus have,

Theorem 1. *Let G be a Sender Receiver game with at least two distinct payoffs to the Receiver and let $\mu(\cdot|t)$ be a rational mediated equilibrium with payoffs to the Receiver greater than the above payoffs. Then, for every $\varepsilon > 0$, there is an extended game $\widetilde{G}$ and a sequential equilibrium s in $\widetilde{G}$ such that the probability of implementing $\mu(\cdot|t)$ is at least $1 - \varepsilon$.*

We now come to the case of games where the set of Sender-Receiver equilibria is a singleton. Here the analysis depends on that of Aumann and Hart [1]. Unfortunately, the characterisation is only of the Bayesian Nash equilibria, and as our result depends on theirs, our theorem will also be correspondingly weaker. Let G be a Sender-Receiver game and p be the distribution of types. Let $\widehat{G}$ be the set of Sender-Receiver game with the property that the fiber over p in the Receiver's payoff in dispan($\mathcal{E}$)[1] is not a singleton. Then we have

Theorem 2. *Let $\widehat{G}$ be defined as above. Then for every mediated equilibrium $\mu(\cdot|t)$ with payoff for the Receiver greater than that available in dispan($\mathcal{E}$), and for all $\varepsilon > 0$, there is an extended game $\widetilde{G}$ with a Bayesian Nash equilibrium s such the probability of implementing $\mu(\cdot|t)$ is at least $1 - \varepsilon$.*

We have thus demonstrated that the lack of commitment in applications is a hurdle that can be overcome virtually and provide a mechanism for achieving this end.

References

[1] Aumann, Robert J. and Sergiu Hart (2002). "Long Cheap Talk," forthcoming in *Econometrica.*

[2] Ben-Porath, Elchanan (1998). "Communication without Mediation: Expanding the Set of Equilibrium Outcomes by "Cheap" Pre-play Procedures," *Journal of Economic Theory, 80,* 108-122.

[3] Bester, Helmut and Rloand Strausz (2001). "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica, 69(4),* 1077-1098.

[4] Myerson, Roger B. (1991). *Game Theory: Analysis of Conflict,* Harvard University Press, Cambridge, Massachusetts and London, England.

[1] The diconvex hull of the graph of equilibrium payoffs of G over p.