DEFINABLE AND CONTRACTIBLE CONTRACTS

MICHAEL PETERS AND BALÁZS SZENTES

Abstract. This paper analyzes a normal form game in which actions as well as contracts are contractible. The contracts are required to be representable in a formal language. We prove a folk-theorem for games with and without privately informed agents. This is accomplished by constructing contracts which are definable functions of the Godel code of every other player's contract. We use this to illustrate the 'meet the competition' argument from Industrial Organization and the 'principle of reciprocity' from Trade and Public Finance.

1. Self Referential Strategies and Reciprocity in Static Games

The idea that players in a game might simultaneously commit themselves to react to their competitors actions is heuristically compelling. The best known expression of this idea is well known in the industrial organization literature (e.g. [8]) as the 'meet the competition' clause. A similar idea appears in trade theory as the principle of *reciprocity* ([1]). This takes the form of trade agreements like GATT that require countries to match tariff cuts by other countries. Finally, tax treaties sometimes have this flavor - for example, out of state residents who work in Pennsylvania are exempt from Pennsylvania tax as long as they live in a state that has a reciprocal agreement that exempts out of state residents (presumably from Pennsylvania) from state taxes.[1]

One way to model reciprocity is to embed it in a dynamic game. For example the tit for tat strategy in the repeated prisoners' dilemma makes each player's action depend on the action of the other. The 'meet the competition' argument could be supported formally by having one firm acts as a Stackleberg leader, offering a contract that commits it to an action that depends explicitly on the action of the second mover. Tax reciprocity could again be accomplished by embedding the problem in a repeated game in which states keep lists of other states whom they consider to have an appropriate tax treaty, deleting a state from the list if they observe some kind of bad behavior. Our interest here is whether this same kind of reciprocal behavior could be modelled in a completely static game.

To illustrate the problem, and our solution, focus first on the meet the competition argument. The Stackleberg leader, call it firm A, offers to sell at a very high price provided its competitor, firm B, also offers that high price in the second round. If B in the second round offers any price below the highest price, A commits itself to sell at marginal cost. If B believes this commitment, then his best reply is to set the highest price. If the firms move simultaneously, then the logic of the argument becomes clouded. A could certainly write a contract that commits it to a high price

[1] http://www.revenue.state.pa.us/revenue/cwp/view.asp?A=238&Q=244681

if B sets the same high price. However suppose that B's strategy is simply to set this high price and that for some reason this is a best reply to A's contract. Then A should deviate and simply undercut firm B. To support the high price outcome, firm B would have to offer a contract similar to A's in order to prevent A's deviation. A naive argument would suggest that B should simply offer the same contract as A, a high price if A sets a high price, and marginal cost otherwise. Casually, two outcomes seem consistent with these contracts - both firms price at marginal cost or both firms set the high price. This seems to violate a fairly fundamental property of game theory which is that for each pair of actions (contracts in this case), there is a unique payoff to every player.[2] More to the point, A's contract doesn't actually say what A would do if B offers a contract that promises to set a high price unless A sets a lower price, etc. The specification of the problem itself seems to be ambiguous about payoffs.

The reciprocal tax agreement also nicely illustrates the difficulty in a static game. State A wants to exempt residents of state B from state taxes provided B exempts residents of state A from taxes. To write the law A exempts residents from any state that has a 'reciprocal' agreement with state A. The question is what exactly is a 'reciprocal' agreement. It is clear enough what the intention is - create a situation in which both states take the mutually beneficial action of exempting one another in a way that eliminates any incentive for either of them to deviate. As mentioned above, it isn't enough to assume that state B unconditionally exempts residents of state A from tax because A would not longer have any incentive to exempt state B. State B has to have a law like the law in state A, in other words, a reciprocal agreement.

It seems that to resolve this kind of problem one needs to define the term 'reciprocal contract' as follows:

$$\text{reciprocal contract} \equiv \begin{cases} \text{exempt if the other state offers a reciprocal contract,} \\ \text{don't otherwise} \end{cases}$$

This kind of definition is familiar from the Bellman equation in dynamic programming where the value function is defined in a self referential way. It is tempting to model this in the following naive way: start by defining a collection of contracts that seem economically sensible. For example, it is reasonable that a state could write a contract that simply fixes any tax rate independent of what the other states do. Let $\overline{C}$ be the set of contracts that simply fix some unconditional tax rate. Append to this set of feasible contracts the reciprocal contract, call it r, defined above. Now model the set of feasible contracts as $\overline{C} \cup \{r\}$. The reciprocal contract above is just r, while 'otherwise' means any contract with a fixed tax rate. Define a normal form game in which the strategies are $\overline{C} \cup \{r\}$ and declare the outcome if both states offer r to be $(\text{exempt}, \text{exempt})$. Voila, there is an equilibrium in which the states mutually exempt (assuming they jointly want to).

[2]One paper that allows multiple payoffs to be associated with each array of actions is [9] who use this approach to support equilibrium when it might not otherwise exist.

We would argue that this is unsatisfactory for a number of reasons. First, it is undesirable to restrict the set of feasible conracts in order to support the outcome you are looking for. The approach described above amount to little more than saying that r is the only feasible contract, then claiming it is an equilibrium for both states to offer r. A more satisfactory approach is to define a set of actions that seem economically meaningful, then to allow the broadest set of contracts possible. In the same manner that the value function emerges endogenously from the economic environment, the reciprocal contract should be derived from economic fundamentals.

One complication that makes this problem conceptually more difficult than the Bellman problem is that it isn't clear what the appropriate set of feasible contracts should be. Clearly states can go further than simple unconditional tax rates. Existing laws do allow them to make things contingent on laws in other states, as the reciprocal tax agreement illustrates. One contribution of our approach is to provide a potential framework for thinking about this set of feasible contracts. As we describe in more detail below, definable functions can all be written as finite sentences in a formal language. In a heuristic sense, this seems exactly what the set of contracts should look like. Of course, in the way it is used here, definability is very abstract. Yet it is abstract in the way that direct mechanisms are abstract - it captures in a formal way 'indirect' contracts that look much more familiar.

Second, the approach described above misses the essence of reciprocity which is the infinite regress involved in self referential objects. A contract that makes formal sense is the following:

$$\mathrm{C} = \begin{cases} \text{exempt if other State exempts any State who exempts any State who exempts...} \\ \text{don't otherwise} \end{cases}$$

where the statement in the top line is repeated ad infinitum. Arguably, the contract C is a reciprocal contract since it would exempt any State offering a reciprocal contract. Yet it simply isn't feasible under the naive description given above.

Of course, in the spirit of the ad hoc approach above, we could try to add the contract C to r and $\overline{C}$. This approach breaks down once the game becomes asymmetric. For example, if State A is supposed to exempt, while state B is supposed to take some other action, say 'partly exempt', then to support the right outcome, the contracts should look something like the following:

$$\text{reciprocal contract}_A \equiv \begin{cases} \text{exempt if other State offers reciprocal contract}_B \\ \text{don't exempt otherwise} \end{cases}$$

and

$$\text{reciprocal contract}_B \equiv \begin{cases} \text{partially exempt if other State offers reciprocal contract}_A \\ \text{don't exempt otherwise} \end{cases}$$

Now the contracts are not directly self referential, as is the Bellman equation, instead they are cross referential. A single self referential or reciprocal contract simply doesn't go far enough. Furthermore, the contracts above use a blanket punishment for deviations. Desirable or interesting equilibrium allocations may not look like this. For example, in a Bayesian game between the states, State A might want to do something different for each different thing that state B might do. This might arise if the action and contract chosen by B convey some information to A that affects A's most desirable action.

In this paper, we offer a formalism that provides a way to think about self-referentiality and reciprocity. Suppose there are N players in a normal form game in which each player has a countable number of actions. Endow players with a formal language that they can use to write contracts and think of the set of feasible contracts as the set of finite sequences of words in this formal language. It is well known that there are bijections from the set of finite texts into $\mathbb{N}$. One such a mapping is called the *Godel Coding.* Provided the language includes all the natural numbers and the usual arithmetic operations, it is possible for players to write contracts that are *definable* functions from $\mathbb{N}^{N-1}$ into that player's action space. Since definable functions can be written as finite sequences of words in the language, they have Godel codes associated with them. Hence we could interpret the definable functions as contracts that make the players action depend on the Godel code of the other player's contract.

To make the argument easier to relate to conventional contract theory, we assume below that the strategy space for each player is the set of definable functions from $\mathbb{N}^{N-1}$ into the player's action space. Implicitly, this approach makes it possible for players to offer any finite text as a contract. We defer discussion of this point to later in the paper. Every definable function can be associated with a unique integer, and conversely if the integer n is associated with a definable function, then it is associated with a unique definable function. Now for each array of functions chosen by the players, compute the Godel Code of each such function. Fit the codes of the other players' strategies into each player's strategy to determine a unique action for every player. Then use the payoffs associated with those actions to define unique payoffs associated with every array of strategies chosen by the players.

Our objective is to try to characterize the set of equilibria of this game. To see how it works, we might as well restrict attention to a two player prisoner's dilemma. Call the players 1 and 2, and the actions C and D with the usual payoff structure in which D is a dominant strategy and both players are strictly better off if they both play C than they are if they both play D. A strategy c for a player is a definable function from $\mathbb{N}$ to $\{C, D\}$. One obvious equilibrium of this game occurs when both players use a strategy that chooses action D no matter what the Godel code of the other player's strategy.

Every definable strategy has a Godel code. Let $[c]$ denote the Godel code of the strategy c and refer to $[c]$ as the 'encoding' of c. Since the Godel coding is an injection from the set of definable strategies to the set of integers. For any pair of strategies c_1 and c_2, the action (C or D) taken by player 1 is $c_1([c_2])$ and similarly for player 2. Since every pair of actions determines a payoff, this procedure associates a unique payoff with every pair of strategies.

There are many things that aren't definable strategies that also have Godel codes. We want to make use of some of these other things. In particular, we want to use definable strategies with *free variables.* For example, there is a subclass of definable strategies for player 1 defined parametrically by

$$\gamma_x(n) = \begin{cases} C & n = x, \\ D & \text{otherwise.} \end{cases}$$

This is simply a definable strategy with a *free variable* x, where x is the target code of the other player's strategy that will trigger the cooperative action. Definable strategies with free variables are also definable, and so they too have Godel codes. The strategy with free variable that we want is a slight modification of the one above, in particular

$$c_x(n) = \begin{cases} C & n = \left[\langle x\rangle^{(x)}\right], \\ D & \text{otherwise.} \end{cases} \tag{1}$$

The mapping $< x >^{(x)}$is the composition of two functions. First, the function $\langle x\rangle$ is the inverse operation to the Godel coding. That is, $< n >$ is the text whose Godel code is n. Second, if ϕ is a text with one free variable, then $\phi^{(n)}$ is the same text where the value of the free variable is set to be n. Hence, if n is a Godel code of a definable strategy with one free variable, then $< n >^{(n)}$ is itself a definable strategy (without a free variable). $\left[\langle n\rangle^{(n)}\right]$ is just the Godel code of whatever this definable strategy happens to be. Notice that in this case, $[\langle x\rangle(x)]$ won't be equal to x since a definable strategy must have a different Godel code from a definable strategy with one free variable because of the fact that the Godel coding is injective.

We want to define a strategy by fixing a value for x in (1). In particular, the value of x we are interested in is $[c_x]$. Since $[c_x]$ is the Godel code of a strategy with a free variable, the right hand side of (1) requires that we decode $[c_x]$ to get c_x, then fix x at $[c_x]$ to get the contract $c_{[c_x]}$. Putting all this together gives

$$c_{[c_x]}(n) = \begin{cases} C & n = \left[c_{[c_x]}\right] \\ D & \text{otherwise} \end{cases}$$

So

$$c_{[c_x]}([c_2]) = \begin{cases} C & [c_2] = \left[c_{[c_x]}\right] \\ D & \text{otherwise} \end{cases}$$

is a the 'reciprocal' or self-referential contract mentioned above. Now we simply need to verify what happens when both players use strategy $c_{[c_x]}$.

If player 2 uses strategy $c_{[c_x]}$, then $[c_2] = \left[c_{[c_x]}\right]$, which evidently triggers the cooperative action by player 1. The same argument applies for player 2. Player 2 can deviate to any alternative definable strategy c' that she likes. Since every definable strategy has a Godel code, the reaction of player 1, and consequently both players payoffs are well defined. As the Godel coding is injective, $c' \neq c_{[c_x]}$ implies the Godel code of c' is not equal to $\left[c_{[c_x]}\right]$, and the deviation by 2 induces 1 to respond by switching from C to D.

Notice that this argument makes use of an encoding of the strategy with free variable c_x, which isn't a definable strategy. One might have expected the target code number to be associated with a strategy instead of a strategy with a free variable. For example, it seems that to enforce cooperation there needs to be a definable strategy c^* with encoding $[c^*] = n^*$ such that

$$c^* = \begin{cases} C & [c_2] = n^* \\ D & \text{otherwise} \end{cases}$$

Of course, for arbitrary n^* it will be false that $[c_{n^*}] = n^*$. This leads to a fixed point problem that, in fact, does not have a solution in general. More generally, one could try to construct a self-referential contract by finding a fixed point of the the following problem. For each n, consider

$$c_n([c_2]) = \begin{cases} C & \text{if } [c_2] = g(n), \\ D & \text{otherwise,} \end{cases}$$

where g is a definable function. If there exists an n^* such that $[c_{n^*}] = g(n^*)$, then c_{n^*} is obviously a self-referential contract. Indeed, what we did above is that we chose $g(n)$ to be $[< n >^{(n)}]$ and showed that $n^* = [c_x]$ is a corresponding fixed point.

To see how the strategy with free variable c_x works, recall the reciprocal tax agreement

$$\text{reciprocal contract} \equiv \begin{cases} \text{exempt} & \text{other State offers reciprocal contract} \\ \text{don't exempt} & \text{otherwise} \end{cases}$$

and its recursive counterpart

$$\text{C} = \begin{cases} \text{exempt if other State exempts any State who exempts any State who exempts...} \\ \text{don't otherwise} \end{cases}$$

The 'reciprocal contract' is $c_{[c_x]}$ and the statement "other state offers reciprocal contract" is $[c_2] = \left[c_{[c_x]}\right]$.

State A wants to exempt any state whose law fullfills a condition. For example, if the condition it is looking for is that the other state simply exempts State S, then it would compute the Godel code $n_0 = [\forall n; \overline{c}(n) = C]$ then use the strategy

$$c_{n_0} = \begin{cases} C & [c_2] = n_0 \\ D & \text{otherwise} \end{cases}$$

If it does that, then it can't be an equilibrium as explained above. So what it needs to do is to exempt any State whose law fullfills a condition that exempts any state whose law fullfills a condition. For example, if it wanted to exempt State B if and only if State B's law exempts state A if and only if State A unconditionally exempts state B, then it would adopt the strategy $c_{\left[c_{n_0}\right]}$, and so on.

This is where the particular structure of the contract c_x comes into play. Recall that

$$c_x(n) = \begin{cases} C & n = \left[\langle x \rangle^{(x)}\right], \\ D & \text{otherwise.} \end{cases}$$

It specifies exemption if and only if a condition is fullfilled, but it doesn't seem to specify what the condition is. However, it does require that whatever the condition x is, if x in turn depends on a condition, then the condition that it depends on must be the same as the condition itself. To see if x depends on a condition, we first decode it and find the statement $\langle x \rangle$ that the integer x corresponds to. Then if it depends on some condition, we require that that condition be x itself, which is the meaning of $\langle x \rangle^{(x)}$. So now we can do the infinite regress. State A adopts a law that exempts state B if and only if the Godel code of State B's law is $\left[c_{[c_x]}\right]$. This means that state B's law must be $c_{[c_x]}$, or that B exempt A if and only if the Godel code of State A's law is $\left[c_{[c_x]}\right]$, i.e., the same condition that A requires.

2. Literature

The approach we develop here is not the only way to sustain cooperation without repetition. An alternative logic has been developed in the theory of common agency ([6] or [7]) in which punishments are carried out through an agent. In problems in which principals interact through many agents, this logic can be used to prove 'folk theorems'. The argument appears in a recent paper by [5] and more generally in the working paper by [10]. The basic logic in the latter paper works as follows - each principle offers a contract with a message space consisting of all the actions that he could take. He then asks the agents to tell him what to do. If all but one of the agents who interact with him names the same action, the principle takes that action, otherwise he takes some default action. It pays the agents to agree with everyone else, because they expect everyone to agree and accomplish nothing by deviating. If all principals offer this contract, then every agent makes the same report to any given principle telling the principal to take some action that, along with the instructions given to the other principals, provides the principle a payoff at least as large as his minmax payoff. If any principle deviates and offers any alternative mapping from messages to outcomes, the agents instruct the other principals to minmax the deviator.

In a very specialized environment, [5] makes an even simpler argument. Working in an environment in which agents take actions on behalf of principles (as in, for example, [3]), Katz imposes enough quasi-linearity and separability such that for any fixed action, the principals can offer the agent a contract such that the optimal effort for the agent under that contract implements the desired action and provides the agent exactly his reservation payoff. Now take any collection of actions for the principles that provides each principal at least his or her minmax payoff. Each principal then offers the agent a contract with a binary message. If the agent sends the message 1, the principle offers the agent a contract that implements his part of the collusive outcome. If the agent sends the message zero, the principle offers a contract that implements the action that minmaxes the other principal. Each agent sends his or her principal the message 1 as long as his principle offers this contract - deviations cause agents to send message zero, and lead to punishment.

One objection to contracts like this, and the folk theorems that they generate, is that they rely heavily on agents coordinating their messages. Yamashita's paper shows this in the most striking way. Not only must agents coordinate on the same message, but the must also coordinate on the message that the principals believe is driving the game before they offer their contracts. Not only is such a strong reliance on equilibrium selection sensitive to common knowledge assumptions, it is also very sensitive to possible collusion or unmodelled communication between agents. All the arguments in this literature basically allow agents to tell principals what they should do. Katz's argument deals with a somewhat simpler environment in which a single agent is hired to carry out an action on the part of the principal. The environment satisfies the assumptions in [4], so that coordinated actions by the principals can be carried out by having each principal offer the agent a menu of options, as they would in common agency. Agents punish deviations by changing the choice they make from this menu. Since agents must be indifferent between the choices in the menu to make this work, the environment is very restricted. Indifference again means that principals must rely on agents choosing appropriately among actions to which they are indifferent.

In this paper we show how to support a folk theorem without the help of agents. This is done by making contracts explicitly refer to one another. This cross referencing feature exploits the concept of definability. Definable contracts can be encoded as integers, so contracts can refer to contracts by referring to the integers associated with them. After all is said and done, the actions the principals take are a joint consequence of the contracts that are offered by all the principals. The implication of this is that contracts themselves can be written in such a way that deviations by one principal trigger responses by other principals without any ambiguity about outcomes. This method can be used to support a folk theorem.

We show how to support a folk theorem in a static normal form game in which there are no agents. The following section adds the agents back and shows how to use the concept of definability to provide a kind of revelation principle for competing mechanism problems, similar to that offered by [2]. In our formulation, the agents are not responsible for instructing the principals to punish. There is no need for them to do that since the contracts principals offer already condition on one another.

These theorems come at a cost - contracts must be definable. In the final section of the paper, we illustrate how definability restricts the physical environment.

3. The Language and the Gödel Coding

We consider a formal language, which is sufficiently rich to allow its user to state propositions in arithmetic. Furthermore, the set of statements in this language is closed under the finite applications of the Boolean operations: $\neg$, $\vee$, and $\wedge$. This implies that one can express, for example, the following statement:

$$\forall n, x, y, z \left\{ \left[(n \geq 3) \vee (x \neq 0) \vee (y \neq 0) \vee (z \neq 0) \right] \rightarrow (x^n + y^n \neq z^n) \right] \right\}.$$

In addition, one can also express statements in the language that involve any finite number of free variables. For example, "x is a prime number" is a statement in the language. The symbol x is a free variable in the statement. Another example for a predicate that has one free variable is "$x < 4$." One can substitute any integer into x and then the predicate is either true or false. This particular one is true if $x = 0, 1, 2, 3$ and false otherwise.

Let $\mathfrak{L}$ be the set of all formulas of the formal language. Each of its element is a finite string of symbols. It is well known that one can construct a one-to-one function $\mathfrak{L} \rightarrow \mathbb{N}$. Let $[\varphi]$ be the value of this function at $\varphi \in \mathfrak{L}$, and call it the Gödel Code of the text φ.

Definition 3.1. The function $f : \mathbb{N}^k \rightarrow 2^{\mathbb{N}}$ is said to be *definable* if there exists a first-order predicate ϕ in $k+1$ free variables such that $b \in f(a_1, ..., a_k)$ if and only if $\phi(a_1, ..., a_k, b)$ is true.

In the definition, the mapping f is a correspondence from $\mathbb{N}^k$ to $\mathbb{N}$. Of course, if $f(n)$ is a singleton for all $n \in \mathbb{N}^k$, then f is a function. We illustrate this definition with an example.

Example. Consider the following function defined on $\mathbb{N}$:

$$f(a) = \begin{cases} 0 & \text{if } a \text{ is an even number,} \\ 1 & \text{if } a \text{ is an odd number.} \end{cases}$$

We show that this function is definable by constructing the corresponding predicate ϕ.

$$\phi(x, y) \equiv \{\{y = 1\} \wedge \{y = 0\}\} \vee \{\exists z : 2z = y + x\}.$$

Notice that ϕ indeed has two free variables. (The variable z is not free because there is a quantifier front of it.) The first part of ϕ states that y is either one or zero. The second part says that $x + y$ is divisible by two. Notice that $f(a) = 0$ if and only if $\phi(a, 0)$ is true. To see this, first notice that $\phi(a, b)$ is false whenever $b \notin \{0, 1\}$. (This is because the first part of ϕ requires b to be zero or one.) If $b = 0$ then $\phi(a, 0)$ is indeed true. If $b = 1$, then the second part of ϕ becomes false because $a + b$ is an odd number.

4. A Normal Form Contracting Game

Suppose there are m players. Each player has a countable action space $A_i = \left\{a_1^i, a_2^i, ...\right\}$.[3] The payoff of Player i is $u_i(a_1, \ldots, a_m)$ if the action taken by player j is a_j. We use the conventional notation that $u_i(a_i, a_{-i})$ is the payoff to player i if he takes action a_i while the other players take action a_{-i}. The game played we consider has two stages. At stage one, player submit *contracts* simultaneously which are definable correspondences from $\mathbb{N}^m$ to $2^{\mathbb{N}}$, where 'definable' is to be understood in the sense of Definition 3.1. At stage two, players take actions simultaneously from a subset of their actions spaces. If at stage one player j submitted contract c_j $(j = 1, ..., m)$,

[3]One notable example is to imagine a underlying finite game in which pure actions are Y_i and A_i represents a finite approximation of the set of mixtures on Y_i.

then player i can only take action a_k^i at stage two if $k \in c_i\left([c_1], ..., [c_m]\right)$. We restrict attention to pure-strategy subgame perfect equilibra of this game.

Our objective is to prove a folk theorem for this contracting game. The lowest payoff for any player in any pure strategy equilibrium of the game in which players choose actions from A is

$$u_i^* = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i\left(a_i, a_{-i}\right),$$

Let a_j^* be any one of the actions that j uses to attain his minmax payoff. Let us fix an action $a_{j_i}^i$ for player i, such that,

$$\left(a_{j_1}^1, ..., a_{j_m}^m\right) \in \arg\min_{a_{-i}} u_j\left(a_j^*, a_{-j}\right).$$

That is, $a_{j_i}^i$ is the action that player i uses to punish player j. In addition, define $i_i = 1$ for all $i \in \{1, ..., m\}$.

Theorem 4.1. *Let $\left\{a_{k_1}^1, \ldots, a_{k_m}^m\right\}$ be any array of actions. These actions are supportable as an equilibrium outcome in the contracting game with pure strategy SPNE if and only if $u_i(a) \geq u_i^*$ for each i.*

Before we proceed with the proof of the theorem, we recall two notations from the introduction. First, if $n \in \mathbb{N}$ then $< n >$ denotes the text whose Gödel code is n. That is, $[< n >] = n$. Second, for any text φ, let $\varphi^{(n_1, \ldots, n_k)}$ denote the statement where if the letter x_i stands for a free variable in φ then x_i is evaluated at n_i in φ for $i = 1, ..., n$. For example, if φ is $x > y$ and $n = 2$, then $\varphi^{(n)}$ is $2 > y$. Consider now the following text in one free variable: $< x >^{(x)}$. One can evaluate this statement at any integer. Since the Godel coding was a bijection $< n >$ is a text for each $n \in \mathbb{N}$. In addition, $\varphi^{(n)}$ is defined for all φ and n. In addition, it is a well-known result in Mathematical Logic, that if $f(n) = \left[< n >^{(n)}\right]$, then f is a definable function.

Proof. First, we prove the only if part. Fix an equilibrium in the contracting game. Let c_j denote the equilibrium contract submitted by player j ($j = 1, ..., m$) and let u_i denote player i's equilibrium payoff. Notice, that player i can always offer a contract that does not restrict his action space. That is, he can offer $\overline{c} : \mathbb{N}^m \to \mathbb{N}$, such that $\overline{c}(n_1, ..., n_m) = \mathbb{N}$ for all $(n_1, ..., n_m) \in \mathbb{N}^m$. The contract $\overline{c}$ is obviously definable.[4] We show that if $u_i < u_i^*$, player i is strictly better off by offering $\overline{c}$ instead of c_i. Let $\widetilde{c}_j = c_j$ if $j \neq i$ and $\widetilde{c}_i = \overline{c}$. Let $\widetilde{A}_j = \left\{a_k^j : k \in \widetilde{c}_j\left([\widetilde{c}_j], ..., [\widetilde{c}_j]\right)\right\}$. That is, $\widetilde{A}_j$ is the action space of player j in the subgame generated by the contract profile $(\widetilde{c}_1, ..., \widetilde{c}_m)$. Also notice that $\widetilde{A_i} = A_i$. The payoff of player i in this subgame is weakly larger than

$$\min_{a_{-i} \in \widetilde{A}_{-i}} \max_{a_i \in A_i} u_i\left(a_i, a_{-i}\right) \geq \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i\left(a_i, a_{-i}\right).$$

The weak inequality follows from $\widetilde{A}_j \subseteq A_j$ for all j. Therefore, player i can always achieve his pure minmax value by offering the contract $\overline{c}$ at stage one.

[4] For example, the predicate

$$\{x_1 = x_1\} \wedge ... \wedge \{x_m = x_m\} \wedge \{y = y\}$$

defines $\overline{c}$. That is, for all $y \in \mathbb{N}$ the predicate is true no matter how the free variables are evaluated.

For the if part, consider the following contract of Player i, $c^i_{x_i,x_{-i}}$, in m free variables:

$$c^i_{x_1,\ldots,x_m}\left(\left\{\left[c^j\right]_{j\neq i}\right\}\right)=$$

$$\left\{\begin{array}{ll} k_i & \text{if } |\left\{k:\left[< x_k >^{(x_1,\ldots,x_m)}\right]\neq\left[c^k\right]\ \right\}|\neq 1, \\ j_i & \text{if } \left\{k:\left[< x_k >^{(x_1,\ldots,x_n)}\right]\neq\left[c_k\right]\ \right\}=\{j\} \end{array}\right. \tag{2}$$

This contract with free variables is a definable function with free variables from $\mathbb{N}^{m-1}$ to $\mathbb{N}$ as long as the actions are replaced with their indices.

The expression (2) is not a contract, but rather a contract with free variables. Each such expression has a Godel code, so let $\gamma_i=\left[c^i_{x_1,\ldots,x_m}\right]$. The functions $\left\{c^i_{\gamma_1,\ldots,\gamma_m}\right\}_i$ have no free variables, so they constitute a set of contracts. We will now show that $\left\{c^i_{\gamma_1,\ldots,\gamma_m}\right\}_{i=1}^m$ constitutes an equilibrium profile of contracts which support the outcome $\left\{a^1_{k_1},\ldots,a^m_{k_m}\right\}$. First observe what happens when all players use contract $c^i_{\gamma_1,\ldots,\gamma_m}$. Notice that

$$c^i_{\gamma_1,\ldots,\gamma_m}\left(\left\{\left[c^j\right]_{j\neq i}\right\}\right)=\left\{\begin{array}{ll} k_i & \text{if } |\left\{k:\left[< \gamma_k >^{(\gamma_1,\ldots,\gamma_m)}\right]\neq\left[c_k\right]\ \right\}|\neq 1, \\ j_i & \text{if } \left\{k:\left[< \gamma_k >^{(\gamma_1,\ldots,\gamma_m)}\right]\neq\left[c_k\right]\ \right\}=\{j\}\,. \end{array}\right.$$

Player i needs to check whether the Godel code of$< \gamma_k >^{(\gamma_1,\cdots,\gamma_m)}$ is equal to the Godel code of $c^k_{\gamma_1,\ldots,\gamma_m}$. The integer γ_k is the Godel code of the contract with free variable $c^i_{x_1,\ldots,x_m}$. Player i's contract says to take this contract with free variable, fix the free variables at $\gamma_1,...,\gamma_m$ (which gives the contract $c^k_{\gamma_1,\ldots,\gamma_m}$), then evaluate its Godel code. This is what is to be compared with the Godel code of the contract offered by k. Of course, these are the same. Since this is the case for all $m-1$ of the other players, player i ends up taking action a_i. So these contracts support the outcome we want if everyone uses them.

Player j can deviate to any definable contract mapping $\mathbb{N}$ into $\mathbb{N}$. However, any such contract will have a different Godel code, and so will induce the punishment $\left\{a^i_{j_i}\right\}_{i\neq j}$ from the other players. Since $u_j(a)\geq u^*_j$ this deviation will be unprofitable. □

Everything about this theorem involves pure strategies. This imposes limits on its application. For example, the game of matching pennies (with payoffs 1 and -1) has *pure strategy* minmax values equal to 1 for both players, and there is no array of pure strategies that supports payoffs at least as high as the pure strategy minmax value for both players. There is no contract equilibrium with a pure selection in this game, indeed there is no pure selection at all. Dropping the restriction to a pure selection allows us to support a (unique in the allocational sense) contract equilibrium in which both players reserve the right to choose their actions non-cooperatively no matter what the other's contract.

The contract equilibrium as we have defined it requires pure strategies in the space of contracts. There is also something that looks at first glance like a 'mixed' equilibrium in contract space in which each player offers each of the degenerate contracts with probability 1/2. There are two complications with showing that this is an equilibrium. The first is that, in general, we have not been able to rule out the possibility that contracts can be written in such a way that one player can contract directly on the actions of the other. For example, when mixing over degenerate contracts in matching pennies, it is conceivable that one player might be able to deviate to a contract that conditions on the outcome of the other players randomization. It is this fact that prevents us from providing a folk theorem for contract games in which the contracts have to specify actions instead of sets of action. In such contract games, we cannot rule out the possibility that contract equilibrium support payoffs below minmax.

The second problem is that the set of definable contracts doesn't necessarily have any nice structure. For example, it isn't obvious how to specify the set of possible mixtures on this set of functions. Absent this, it is hard to know what all the possible deviations are. As a result it isn't obvious how to prove that mixtures over degenerate contracts consitute an equilibrium.

One might argue that restricting the space of contracts to be definable functions of Godel codes is both arbitrary and unnatural. Indeed, there is no reason for a judge to interpret a contract as a description of a mapping from the Godel codes of the contracts offered by the other players to the actions space of the player. For that matter, the judge might not even know about the Godel coding. It is important to note that the salient feature of definable contracts is that they can be written as texts that use a finite number of words in a formal language. The set of finite texts seems a very natural description of the set of feasible contracts. In fact, from this perspective is seems that *any* reasonable description of the set of feasible contracts should allow any such text.

The complication with such a broad description of the set of contracts is that to properly define a game, one must fully describe the mappings from profiles of texts into payoffs. Many texts will be complete nonsense and some modelling decision has to be taken about how these would translate into actions and payoffs. The contracts that we specify above are definable texts that have two advantages in this regard. First, since every finite text has a Godel code, they tie down the action of the player who offers such a contract even if the other players in the game offer contracts involving texts that make no economic sense. Furthermore, if all players offer contracts from the set we specify, an outcome for every player is uniquely determined. So no matter how ambiguous outcomes are set when players offer non-sense texts as contracts, the equilibrium outcome we describe above will persist.

From the perspective of the judge who has never heard of a Godel code, Theorem 4.1 and the theorems that follow have a kind of normative implication. This is simply that self or cross referential contracts that to involve an infinite regress can nonetheless be unambiguously written using finite texts.

5. Contracting in a Bayesian Environment

This section shows how to extent the result of the previous section to games with incomplete information. To that end, consider the following model. There are n players. Player i's actions space is a finite set denoted by $A^i = \left\{a_1^i, a_2^2, ...\right\}$. The state of the word is randomly drawn from the set Ω. After the state of the world, ω, is realized, each player receives a signal about ω. The signal space of player i is denoted by T^i, which is a finite set. The joint distribution of the state of the world and the signals are common knowledge. The payoff of player i is $u_i\left(a_i, a_{-i}, \omega\right)$.

Our goal is to characterize the set of equilibria of the contracting game (to be defined formally later) in this physical environment. Our strategy is to define a mechanism design problem, and then to show that the set of allocations that can be implemented by these mechanisms are identical to the set of equilibrium outcomes in the contracting game. To this end, we first describe a class of mechanisms.

The Mechanisms.— Consider the following set of four-stage mechanisms. First, players simultaneously decide whether or not to participate. These decisions are only observable by the mechanism designer (MD). Second, those players who decide to participate, send *public* messages to the MD. Player i's message space, M_i, is a countable set. Third, the MD can arbitrarily restrict the action space of those players who have decided to participate at the first stage, as a function of the messages. Finally, players take actions simultaneously. Those players who did not participate can take any action. If a player participated at stage one, she can only take action from the restricted set. (Of course, these actions can depend on the messages sent at stage two.)

We restrict attention to deterministic mechanisms and pure-strategy Bayesian Nash Equilibriua. By standard arguments in mechanism design, without loss of generality, one can assume that each player participates in the mechanism. In addition, one can aslo assume that each player's type space is partitioned, and the messages reported at the second stage by player i are elements of his partition. This is because of the restriction to pure strategies. Finally, one can restrict attention to those mechanism-equilibrium pairs in which players report their type truthfully. That is, at stage two, a player reports the element of his partition which contains his true type.

Before we proceed with the analysis of this problem, we introduce some notations. Let ς^i denote a partition of the type space of player i. Let $R^i\left(\tau^{j_1}, ..., \tau^{j_k}\right) \subset A^i$ denote the restricted action space of player i if players $j_1, ..., j_k$ participate, and the message sent by player i_q is $\tau^{i_q} \in \varsigma^{i_q}$ $(q = 1, ..., k)$. (If player i does not participate at stage one, $i \notin \{j_1, ..., j_k\}$, then $R^i\left(\tau^{j_1}, ..., \tau^{j_k}\right) = A_i$.) Let $s^i_{j_1,...,j_k} : T^i \times \left(\times_{q=1}^k \varsigma^{j_q}\right) \to A^i$, such that $s^i_{j_1,...,j_k}\left(t^i, \tau^{j_1}, ..., \tau^{j_k}\right) \in R^i\left(\tau^{j_1}, ..., \tau^{j_k}\right)$, the strategy of player i at stage four. Let $s^{-i}_{j_1,...,j_k}\left(t^{-i}, \times_{q=1}^k \tau^{j_q}\right)$ denote the $n-1$ dimensional vector whose coordinates are $\left\{s^q_{j_1,...,j_k}\left(t^q, \tau^{j_1}, ..., \tau^{j_k}\right) : q \neq i\right\}$.

An equilibrium is characterized by the following three constraints. In the last stage, players optimally choose their action given the others' strategies, and truthful reports at stage two. That is, for all i, $t^i \in T^i$, and $\times_{q=1}^k \tau^{j_q} \in \times_{q=1}^k \varsigma^{j_q}$:

$$
\begin{aligned}
(3) \qquad & s^i_{j_1,...,j_k}\left(t^i, \times_{q=1}^k \tau^{j_q}\right) \\
= \ & \arg \max_{a \in R^i\left(\tau^{j_1},...,\tau^{j_k}\right)} E_{t^{-i},\varpi}\left(u_i\left(a, s^{-i}_{j_1,...,j_k}\left(t^{-i}, \times_{q=1}^k \tau^{j_q}\right), \varpi\right) : t^i, \times_{q=1}^k \tau^{j_q}\right).
\end{aligned}
$$

In the expectations, $t^{j_q} \in \tau^{j_q}$ is taken into account. Observe that these constraints have to be satisfied only for $k = n$ and $n-1$. (That is, it has to be satisfied if either one or no player deviated in the first stage.)

Second, at stage 2, players truthfully report the element of the partition given the strategies in the last stage, and that everybody else reports truthfully. That is, for all $i \in \{1, ..., n\}$, and $t^i \in T^i$, if $t^i \in \tau^i$, then for all $\tau'^i \in \varsigma^i$:

$$\begin{aligned} (4) \qquad & E_{t^{-i},\varpi}\left(u_i\left(s^i_{1,...,n}\left(t^i, \times^n_{q=1}\tau^q\right), s^{-i}_{1,...,n}\left(t^{-i}, \times^n_{q=1}\tau^q\right), \varpi\right) : t^i, \times^n_{q=1}\tau^q\right) \\ \geq \quad & E_{\tau^{-i}}\left(\max_{a \in R^i(\tau^1,...,\tau'^i,...,\tau^n)} E_{t^{-i},\varpi} u_i\left(a, s^{-i}_{1,...,n}\left(t^{-i}, \times^{i-1}_{q=1}\tau^q \times \tau^i \times^n_{q=i+1}\tau^q\right), \varpi\right) : t^i, \times^n_{\substack{q=1 \\ q \neq i}}\tau^q\right). \end{aligned}$$

Finally, at stage one, players prefer to participate to opting out, given that everybody else participates, truthful reports at the second stage, and the strategies at the final stage. To characterize this constraint, we first compute the payoff of player i if she does not participate. For all t^i and $\times_{q \neq i}\tau^q$, player i's action at stage four solves

$$\max_{a \in A_i} E_{t^{-i},\varpi}\left(u_i\left(a, \left(s^{-i}_{-i}\left(t^{-i}, \times_{q \neq i}\tau^q\right)\right), \varpi\right) : t^i, \times_{q \neq i}\tau^q\right),$$

where $t_j \in \tau^j$ for all $j \neq i$. Let us denote the solution to this problem by $a^i\left(t^i, \tau^{-i}\right)$. Then, player i prefers to participate if and only if

$$\begin{aligned} (5) \qquad & E_{\varpi,t^{-i}}\left(u_i\left(a^i\left(t^i, \tau^{-i}\right), s^{-i}_{-i}\left(t^{-i}, \times_{q \neq i}\tau^q_{j_q}\right), \varpi\right) : t^i\right) \\ \leq \quad & E_{\varpi,t^{-i}}\left(u_i\left(s^i_{1,...,n}\left(t^i, \times^n_{q=1}\tau^q_{j_q}\right), s^{-i}_{1,...,n}\left(t^{-i}, \times^n_{q=1}\tau^q_{j_q}\right), \varpi\right) : t^i\right). \end{aligned}$$

Therefore an allocation that can be implemented by the mechanism is characterzied by the partitions of the type spaces, $\left\{\varsigma^i\right\}^n_{i=1}$, the restrictions of the MD, $\left\{R^i_{j_1,...,j_k} : i, k, j_q = 1, ..., n\right\}$, and the strategies $\left\{s^i_{j_1,...,j_k} : i, k, j_q = 1, ..., n\right\}$ and the three constraints (3), (4), and (5).

The Contracting Game.— Let us now formally defined the contracting game. The game has two stages. In the first stage, players offer contracts simultaneously. A contract is a pair of a definable function and a subset of the type sapce of the player who is offering the contract. That is, a contract can include a statement about the type of the players. The definable function is a mapping from $\mathbb{N}^{2n}$ to subsets of A_i. The first n coordinates of the domain are the Godel codes of the definable functions offered by the players. The other coordinates are the codings of the subsets of the type spaces of the players. For notational convinience, we shall write these functions as if the last n coordinates were the subsets of types instead of their encodings. If player q profferes a definable function c^q, and she τ^q as the subset of his type space, then $c^i\left(\left[c^1\right], ..., \left[c^n\right], \tau_1, ..., \tau_n\right) \subset A^i$ is the subset of the action space of player i pinned down by the contracts. At stage two, players take actions simultaneously from the subsets of their actions spaces which specified by the contracts.

Theorem 1. A deterministic allocation can be implemented by pure strategies in the mechanism described above if and only if it can be implemented by pure strategies in the contracting game.

Proof. First, let us fix a mechanism and an equilibrium of it. Recall, that every equilibrium pinns down a partition of the type space for all i. Let us denote the partitions of the type spaces by $\left\{\varsigma^i\right\}_{i=1}^n$. The restrictions in the mechanism are $\left\{R^i_{j_1,...,j_k} : i,k = 1,...,n, j_i \in \mathbb{N}\right\}$ and the equilibrium strategies are $\left\{s^i_{j_1,...,j_k} : i,k = 1,...,n, j_i \in \mathbb{N}\right\}$. Consider the following contract in n free variables.

$$\begin{aligned} & c^i_{x_1,...,x_n}\left(c^1,...,c^n,\tau^1,...,\tau^n\right) \\ = & \left\{\begin{array}{ll} R^i_{1,...,n}\left(\times^n_{q=1}\tau^q\right) & \text{if } \left|\left\{k :< x_k >^{(x)} \neq c^k \text{ or } \tau^k \notin \varsigma^k\right\}\right| \neq 1, \\ R_{-j}\left(\times_{q\neq j}\tau^q_{j_q}\right) & \text{if } \left\{k :< x_k >^{(x)} \neq c^k \text{ or } \tau^k \notin \varsigma^k\right\} = j. \end{array}\right. \end{aligned}$$

Let γ_i denote the Godel Code of it. The equilibrium contract offered by player i will be: $c^i_{\gamma_1,...,\gamma_n}$. Then

$$\begin{aligned} & c^i_{\gamma_1,...,\gamma_n}\left(c^1,...,c^n,...,\tau^1,...,\tau^n\right) \\ = & \left\{\begin{array}{ll} R^i_{1,...,n}\left(\times^n_{q=1}\tau^q\right) & \text{if } \left|\left\{k :< \gamma_k >^{(\gamma_1,...,\gamma_n)} \neq c^k \text{ or } \tau^k \notin \varsigma^k\right\}\right| \neq 1, \\ R_{-j}\left(\times^n_{q\neq j}\tau^q\right) & \text{if } \left\{k :< \gamma_k >^{(\gamma_1,...,\gamma_n)} \neq c^k \text{ or } \tau^k \notin \varsigma^k\right\} = j. \end{array}\right. \end{aligned}$$

Notice that $< \gamma_q >^{(\gamma_1,...,\gamma_n)} = c^q_{\gamma_1,...,\gamma_n}$. Therefore, the previous contract can be rewritten as

$$\begin{aligned} & c^i_{\gamma_1,...,\gamma_n}\left(c^1,...,c^n,...,\tau^1,...,\tau^n\right) \\ = & \left\{\begin{array}{ll} R^i_{1,...,n}\left(\times^n_{q=1}\tau^q\right) & \text{if } \left|\left\{k : c^k_{\gamma_1,...,\gamma_n} \neq c^k \text{ or } \tau^k \notin \varsigma^k\right\}\right| \neq 1, \\ R_{-j}\left(\times^n_{q\neq j}\tau^q\right) & \text{if } \left\{k : c^k_{\gamma_1,...,\gamma_n} \neq c^k \text{ or } \tau^k \notin \varsigma^k\right\} = j. \end{array}\right. \end{aligned} \tag{6}$$

We shall argue that for player i $(i = 1,...,n)$ with type t^i offering $\left(c^i_{\gamma_1,...,\gamma_n},\tau^i\right)$, where $t^i \in \tau^i \in \varsigma^i$, at the contracting stage and taking action according to $s^i_{j_1,...,j_k}$ in the second stage constitue an equilibrium. Obviously, the strategies $s^i_{j_1,...,j_k}$ are optimal in the second stage by (3). We only have to show that players do not have incentive to deviate at the first stage when offering contracts. Notice that, by the first line of (6), if a player does not deviate, her payoff is the same as in the mechanism. Suppose now that player i with type t^i offers a contract $\left(c^i,\tau^{i'}\right)$ which is different from $\left(c^i_{\gamma_1,...,\gamma_n},\tau^i\right)$, where $t^i \in \tau^i \in \varsigma^i$. We shall consider two cases. Case 1: $c^i \neq c^i_{\gamma_1,...,\gamma_n}$ or $\tau^{i'} \notin P^i$. (That is, player i offers an off-equilibrium contract.) Then, by the second line of (6), the payoff of the deviator cannot exceed the payoff of player i with type t^i in the mechanism if she decided not to participate in the first stage. Since (5) holds, such a deviation is not profitable. Case 2: $c^i = c^i_{\gamma_1,...,\gamma_n}$ and $\tau^{i'} \in \varsigma^i$ but $\tau^i \neq \tau^{i'}$. Then, again by the first line of (6), the deviator's payoff cannot exceed the payoff of playar i with type t^i in the original mechanism if she decided to report $\tau^{i'}$ instead of τ^i in the second stage. Since (4) holds, such a deviation cannot be profitable.

Let us now prove the converse. Fix an equilibrium in the contracting. We shall argue that the same outcome can be implemented by a mechanism. Notice that by offering a contract, a player reveals some information about his type. Since we restricted attention to pure strategies, the equilibrium contracts generate a partition on the type space of each players. That is, an element of a partition of the type space of player i is the largest set of those types who offer the same contract. In the mechanism, players can report elements of these partitions. If each player participates, the restrictions the MD imposes on the action spaces are the same as the ones imposed by the equilibrium contracts. If a single player does not participate, the restrictions are the same as the ones imposed by the equilibrium contracts in case the deviator offered a contract which did not restrict his action space. Since we have to check only the profitability of a unilateral deviation, the rest of the restrictions can be arbitrary. Finally, the second stage strategies of the players in the contracting games are identical to the strategies in the last stage of the mechanism. It is obvious that the equilibrium payoff of a player as well as a payoff of a deviator are identical in the mechanism and the contracting game. $\square$

The next example illustrates that (i) reporting types (instead of a partition elements) and (ii) restricting players to single actions (instead of subsets of their action spaces) cannot be assumed. This example also shows that there are allocations that can be implemented by the standard revelation principle but not with contractible contracts.

Example. Suppos that there are two players 1 and 2. Player 1 has four possible types, $T^1 = \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$. The probability of each type is one fourth. Player 2's type space is degenerate. Actions of player 1 are $\{1, 2\}$. Actions of player 2 are $\left\{a_1, a_2, b_1, b_1, g_{\alpha_1}, g_{\alpha_2}, g_{\beta_1}, g_{\beta_2}\right\}$. Payoffs are defined as follows.

$$\begin{aligned}
u_j(\alpha_i, k, x) &= u_j(\beta_i, k, x) = -1, \text{ if } k \neq i,\ x \neq G_t,\ \ (i, j = 1, 2),\\
u_j(\alpha_i, i, A_i) &= u_j(\beta_i, i, B_i) = 10,\ \ (i, j = 1, 2),\\
u_j(\alpha_i, i, A_l) &= u_j(\beta_i, i, B_l) = 9 \text{ if } l \neq i,\ \ (i, j = 1, 2),\\
u_1(t, i, G_t) &= 0,\ u_2(t, i, G_t) = 15,\ \ i = 1, 2,\ \ t \in T^1,\\
u_j(t, i, G_{t'}) &= 0 \text{ if } t' \neq t,\ \ (i, j = 1, 2).
\end{aligned}$$

The idea of this game is the following. Player one wants match the index of his type with his action. That is, he wants to take i if his type is α_i or β_i $(i = 1, 2)$. The payoffs of the players are high, ten, if player 2 can match player 1's type with his action perfectly. (That is, if he takes action $a_i\,(b_i)$ if $t = \alpha_i\,(\beta_i)$.) Their payoffs are nine if player 2 matches the type of player one imperfectly, that is, he takes action $A_l\,(B_l)$ if the type is $\alpha_k\,(\beta_k)$ $(l \neq k)$. The twist of the game is the following. If player 2 knows the type of player 1, he can guess it, that is, he can take an action from $\left\{G_t : t \in T^1\right\}$. If he guesses right, he gets a payoff of 15 and player 1 gets zero. If he guesses wrong both players receive a payoff of zero.

Consider the following allocation: $f(\alpha_i) = (c_i, A_1)$ and $f(\beta_i) = (c_i, B_1)$ $(i = 1, 2)$. The expected payoff of both players is 9.5. This is because with probability one half, $i = 1$ and player 2 matches the type perfectly. However, with probablity one half $i \neq 1$, and player 2 matches the type imperfectly. The mechanism-equilibrium pair that implements this allocation can be described as follows. The partition elements of the type space are $\{\alpha_1, \alpha_2\}$ and $\{\beta_1, \beta_2\}$. The restriction on the action space of player 2 of the MD is A_1 (B_1) if player 1 reported $\{\alpha_1, \alpha_2\}$ $(\{\beta_1, \beta_2\})$, and G_{α_1} if player 1 did not participate. The restriction on player 1's action space is $\{1, 2\}$ always. The strategy of player 1 is to take action 1 if $t \in \{\alpha_1, \beta_1\}$ and action 2 otherwise. It is easy to show that players prefer to participate, and player 1 truthfully reports the element of the partition. Hence, the allocation f can indeed be implemened by a mechanism.

Notice that the allocation f cannot be implemented by requiring player 1 to report his type. This is because, in that case, player 2 would guess his type. Also notice that the MD cannot restrict player 1 to take a single action in the last stage, because then he would not be able to match the index of his type.

Consider now the following allocation, F. $F(\alpha_i) = (i, a_i)$ and $F(\beta_i) = (i, b_i)$ for $i = 1, 2$. This allocation can obviously be implemented by the Revelation Principle, but not with contractible contracts.

References

[1] Kyle Bagwell and Robert W. Staiger. Reciprocity, non-discrimination and preferential agreements in the multilateral trading system. *European Journal of Political Economy*, 17(2):281–325, June 2001.

[2] L. Epstein and M. Peters. A revelation principle for competing mechanisms. *Journal of Economic Theory*, 88(1):119–161, September 1999.

[3] C Fershtman and K.L. Judd. Equilibrium incentives in oligopoly. *American Economic Review*, 77:927–940, 1987.

[4] Seungjin Han. Menu theorems for bilateral contracting. *Journal of Economic Theory*, 127(1):157–178, November 2006. available at http://ideas.repec.org/a/eee/jetheo/v131y2006i1p157-178.html.

[5] Michael Katz. Observable contracts as commitments: Interdependent contracts and moral hazard. *Journal of Economics and Management Strategy*, 15(3):685–706, September 2006.

[6] David Martimort and Lars Stole. Communications spaces, equilibria sets and the revelation principle under common agency. University of Chicago unublished manuscript, 1998.

[7] Michael Peters. Common agency and the revelation principle. *Econometrica*, 69(5):1349–1372, September 2001.

[8] S.C. Salop. *Practises that Credibly Facilitate Ologopoly Coordination.* Cambridge, MIT Press, 1986.

[9] Leo Simon and W. Zame. Discontinuous games and endogenous sharing rules. *Econometrica*, 58(4):861–872, 1990.

[10] Takuro Yamashita. A revelation principle and a folk theorem without repition in games with multiple principles and agents. manuscript, Stanford University, February 2007.

Very Preliminary Version - 24th May 2008