Supplement to "Reputation without commitment in finitely repeated games"

(*Theoretical Economics*, Vol. 11, No. 1, January 2016, 157–185)

Jonathan Weinstein
Department of Economics, Washington University in St. Louis

Muhamet Yildiz
Department of Economics, Massachusetts Institute of Technology

S.1. Indistinguishability of testable predictions

The strategic equivalence in Section 5 implies that the testable predictions with or without commitment types are nearly indistinguishable. Imagine that an empirical or experimental researcher observes outcomes of games that essentially look like a fixed repeated game, as in g^*, but she does not know the players' beliefs about possible commitments or payoff variations. Using the data, she can obtain an empirical distribution on outcome paths. Because of sampling variation, there is some noise regarding the actual equilibrium distribution of the outcomes. The above strategic equivalence implies that the equilibrium distributions for elaborations with or without commitment types can be arbitrarily close, making it impossible to rule out one model without ruling out the other given the sampling noise.

Toward stating this result formally, let Σ^* be the set of solution concepts that are (1) invariant to the elimination of nonrationalizable plans, (2) invariant to trivial enrichments of the type spaces, and (3) include all solutions generated by the sequential equilibria that satisfy Assumption 1. Given any solution concept $\Sigma \in \Sigma^*$ and any Bayesian game G, a solution σ leads to a probability distribution $\mathbf{z}(\cdot \mid \sigma) \in \Delta(Z)$ on the set Z of outcome paths, such that

$$\mathbf{z}(z \mid \sigma) = \sum_{\tau \in \mathcal{T}} \sum_{s \in S_z} \sigma(s \mid \tau)\pi(\tau) \quad (\forall z \in Z),$$

where $\mathcal{T}$ is the sets of type profiles in G, $S_z = \{s \in S \mid z(s) = z\}$ is the set of profiles of action plans that lead to z, π is the (induced) common prior on $\mathcal{T}$, and $\sigma(s \mid \tau)$ is the probability of action plan s in equilibrium σ when the type profile is τ. A solution concept Σ yields a set

$$\mathcal{Z}(\Sigma, G) = \{\mathbf{z}(\cdot \mid \sigma) \mid \sigma \in \Sigma(G)\}$$

Jonathan Weinstein: j.weinstein@wustl.edu
Muhamet Yildiz: myildiz@mit.edu

DOI: 10.3982/TE1893

of probability distributions on outcome paths. Toward comparing the distance between such sets, we endow the set $2^{\Delta(Z)}$ of such subsets with the Hausdorff metric d, the standard metric for sets.[1] For any $X, Y \in 2^{\Delta(Z)}$,

$$d(X, Y) \le \lambda$$

if and only if for each $x \in X$, there exist $y \in Y$ and $p \in \Delta(Z)$ with $x = (1-\lambda)y + \lambda p$, and for each $y \in Y$, there exist $x \in X$ and $p \in \Delta(Z)$ with $x = (1-\lambda)y + \lambda p$.

Our first corollary states that the set of distributions on the outcome paths are nearly identical with or without commitment types.

Corollary S.1. *For any $\Sigma \in \Sigma^*$, any ε-elaboration G with commitment types, and any $\varepsilon' \in (\varepsilon, 1)$, there exists an ε'-elaboration G' without commitment types such that $d(\mathcal{Z}(\Sigma, G), \mathcal{Z}(\Sigma, G')) \le (\varepsilon' - \varepsilon)/(1-\varepsilon)$.*

Proof. Define $\lambda = (\varepsilon' - \varepsilon)/(1-\varepsilon)$. Consider the ε'-elaboration G' in Propositions 1–4 of our main paper. Recall that any type profile (τ_1, τ_2^*) in G has identical solutions to a type profile $(f(\tau_1), \tau_2^*)$ in G', where $f(\tau_1^*) = \tau_1^*$ and $f(c) = \tau_1^c$. Moreover, $\pi'(f(\tau_1), \tau_2^*) = (1-\lambda)\pi(\tau_1, \tau_2^*)$. Hence, $y \in \mathcal{Z}(\Sigma, G')$ if and only if there exists $\sigma' \in \Sigma(G')$ such that $y = (1-\lambda)x + \lambda p$ for x and p, where

$$x(z) = \sum_{\tau \in \mathcal{T}} \sum_{\{s \in S | z(s) = z\}} \sigma'(s \mid f(\tau_1), \tau_2^*) \pi(f(\tau_1), \tau_2^*)$$

and $p(z) = \sum_{\tau_2 \ne \tau_2^*} \sum_{\{s \in S | z(s) = z\}} \sigma'(s \mid \tau) \pi'(\tau)$. Since the sets of solutions for (τ_1, τ_2^*) and $(f(\tau_1), \tau_2^*)$ are identical, $x \in \mathcal{Z}(\Sigma, G)$, and the converse is also true in that there exists a $\sigma' \in \Sigma(G')$ as above for every $x \in \mathcal{Z}(\Sigma, G)$. □

Suppose that one wants to restrict G' to be an ε-elaboration, so that the prior probabilities of rational types are identical. The results in the reputation literature are often continuous with respect to ε when the set and the relative probability of the commitment types are fixed. In that case, such a restriction would not make a difference, as established in the next corollary.

Corollary S.2. *Consider any ε-elaboration G with commitment types (C, π) and a solution concept $\Sigma \in \Sigma^*$ such that $\Sigma(G^\alpha)$ is continuous with respect to α, where G^α is an $\alpha\varepsilon$-elaboration with commitment types $(C, \pi/\alpha)$ for $\alpha \ge 1$. Then, for any $\lambda > 0$, there exists an ε-elaboration G' without commitment types such that $d(\mathcal{Z}(\Sigma, G), \mathcal{Z}(\Sigma, G')) \le \lambda$.*

Proof. Apply the previous result starting from G^α for some $\alpha > 1$ that is sufficiently close to 1—in particular, where $\alpha\varepsilon \le \lambda(1-\varepsilon) + \varepsilon$—and then apply continuity. □

[1] More specifically, we use the Hausdorff metric induced by the "total variation" metric on $\Delta(Z)$, but since we only use the metric on sets, we will simply define the Hausdorff metric directly.

S.2. Proof of Lemma 3

We first introduce a more general notion of equivalence. Recall that $z(s)$ denotes the outcome of a profile s of action plans. In line with our notation for histories, we will write $z(s)^t$ for the truncation of $z(s)$ at the beginning date t; i.e., if $z(s) = (a^0, a^1, \dots, a^{\bar{t}})$, then $z(s)^t = (a^0, a^1, \dots, a^{t-1})$. Recall also that action plans s_i and s_i' are *equivalent* if $z(s_i, s_{-i}) = z(s_i', s_{-i})$ for all action plans $s_{-i} \in S_{-i}$, i.e., they lead to the same outcome no matter what strategy the other player plays. Note that s_i and s_i' are equivalent if and only if $s_i(h^t) = s_i'(h^t)$ for every history h^t in which i played according to s_i throughout; they may differ only in their prescriptions for histories that they preclude. Hence, in reduced form, action plans can be represented as mappings that map the history of other players' play into own stage-game actions. Similarly, action plans s_i and s_i' are said to be *t-equivalent* if $z(s_i, s_{-i})^t = z(s_i', s_{-i})^t$ for all action plans $s_{-i} \in S_{-i}$, i.e., they lead to the same history up to date t no matter what strategy the other player plays. Because we have a finite horizon $\bar{t}$, equivalence is the same as $\bar{t}+1$ equivalence. Given any two sets X, Y of action plans, we write $X \simeq^t Y$ if for every $x \in X$, there exists $y \in Y$ that is equivalent to x, and for every $y \in Y$, there exists $x \in X$ that is t-equivalent to y. We prove the following more general version of Lemma 3 for t equivalence. Note that the construction in this proof relies on the fact that players do not know their own stage-game payoffs and do not observe them at each stage, but can learn them from other players' actions.

Lemma S.1 (Weinstein and Yildiz 2013). *For any sure-thing compliant action plan s_i and any $t \in T$, there exists a game $\tilde{G} = (N, A, (\tilde{\mathcal{G}}, \tilde{\mathcal{T}}, \tilde{\pi}(\cdot \mid \cdot)))$ with a type $\tau_i^{s_i,t}$ such that $S_i^\infty[\tau_i^{s_i,t} \mid \tilde{G}] \simeq^t \{s_i\}$. (The type space does not necessarily have a common prior.)*

Proof. We will induct on t. When $t = 1$, it suffices to consider a type $\tau^{s_i,t}$ who is certain that in the stage game, $s_i(\varnothing)$ yields payoff 1 while all other actions yield payoff 0. Now fix t, s_i and assume the result is true for all players and for $t-1$. In outline, the type we construct will have payoffs that are completely insensitive to the actions of the other players, but will find those actions informative about his own payoffs. He also will believe that if he ever deviates from s_i, the other players' subsequent actions are uninformative: this ensures that he always chooses the myopically best action.

Formally, let $\hat{H}$ be the set of histories of length $t-1$ in which player i always follows the plan s_i, so that $|\hat{H}| = |A_{-i}|^{t-1}$, where A_{-i} is the set of profiles of static moves for the other players. For each history $h \in \hat{H}$, we construct a pair (τ_{-i}^h, g^h), and our constructed type $\tau^{s_i,t}$ assigns equal weight to each of $|A_{-i}|^{t-1}$ such pairs. Each type τ_{-i}^h is constructed by applying the inductive hypothesis to a plan s_{-i}^h, which plays according to history h as long as i follows s_i and simply repeats the previous move forever if player i deviates. Such plans are sure-thing compliant for the player $-i$ because at every history, the current action is repeated on at least one branch.

To define the payoff functions θ^h for all $h \in \hat{H}$, we will need to define an auxiliary function $f : \tilde{H} \times A_i \to \mathbb{R}$, where $\tilde{H}$ is the set of prefixes of histories in $\hat{H}$. The motive behind the construction is that $f(h, \cdot)$ represents i's expected value of his stage-game payoffs conditional on reaching the history h. The function f is defined iteratively on histories of increasing length. Specifically, define f as follows: Fix $\varepsilon > 0$. Let $f(\varnothing, s_i(\varnothing)) = 1$

and $f(\varnothing, a) = 0$ for all $a \neq s_i(\varnothing)$, where $\varnothing$ is the empty history. Next, assume $f(h, \cdot)$ has been defined and proceed for the relevant one-step continuations of h as follows.

Case 1. If $s_i(h, (s_i(h), a_{-i})) = s_i(h)$ for all a_{-i}, then let $f((h, a), \cdot) = f(h, \cdot)$ for every a.

Case 2. Otherwise, by sure-thing compliance, at least two different actions are prescribed for continuations $(h, (s_i(h), a_{-i}))$ as we vary a_{-i}. For each action $a_i \in A_i$, let $S_{a_i} = \{a_{-i} : s_i(h, (s_i(h), a_{-i})) = a_i\}$ be the set of continuations where a_i is prescribed. Then let

$$f((h, (s_i(h), a_{-i})), a_i) = \begin{cases} f(h, s_i(h)) + \varepsilon & \text{if } a_{-i} \in S_{a_i} \\ \frac{|A_{-i}| f(h, a_i) - |S_{a_i}|(f(h, s_i(h)) + \varepsilon)}{|A_{-i}| - |S_{a_i}|} & \text{if } a_{-i} \notin S_{a_i}, \end{cases}$$

where the last denominator is nonzero by the observation that at least two different actions are prescribed.

These payoffs were chosen to satisfy the constraints

$$f(h, a_i) = \frac{1}{|A_{-i}|} \sum_{a_{-i}} f((h, (s_i(h), a_{-i})), a_i) \tag{S.1}$$

$$f(h, s_i(h)) \geq f(h, a_i) + \varepsilon \quad (\forall h, a_i \neq s_i(h)),$$

as can be verified algebraically.

For each history $h \in \hat{H}$, define the stage-game payoff function $g^h : A \to [0, 1]^n$ by setting $g_i^h(a) = f(h, a_i)$ and $g_j^h(a) = 0$ at each a and $j \neq i$. Define $\tau^{s_i, t}$ as mentioned above, by assigning equal weight to each pair (τ_{-i}^h, θ^h).

We claim that under rationalizable play, from the perspective of type $\tau^{s_i, t}$, when he has followed s_i and reaches history $h \in \tilde{H}$, $f(h, \cdot)$ is his expected value of the stage-game payoff g_i. We show this by backward induction on the length of histories. When a history $h \in \hat{H}$ is reached, player i knows (assuming rationalizable play) the opposing types must be τ_{-i}^h and thus the stage-game payoff function must be g^h, which is the desired result for this case. Suppose the claim is true for all histories in $\tilde{H}$ of length M. Note that type $\tau^{s_i, t}$ puts equal weight on all sequences of play for his opponent. Therefore, for a history $h \in \tilde{H}$ of length $M - 1$, the expected payoffs are given by the right-hand side of (S.1), which proves the claim.

Note also that if he follows s_i through period t, player i always learns his true payoff. Let $\bar{s}_i$ be the plan that follows s_i through period t and then plays the known optimal action from period $t + 1$ onward. We claim that $\bar{s}_i$ strictly outperforms any plan that deviates by period t. The intuitive argument is as follows. Because type $\tau^{s_i, t}$ has stage-game payoffs that are insensitive to the other players' moves, he only has two possible incentives at each date: the myopic goal of maximizing his average stage-game payoffs at the current date and the desire to receive further information about his payoffs. The former goal is strictly satisfied by the move prescribed by $\bar{s}_i$, and the latter is at least weakly satisfied by this move, since after a deviation he receives no further information.

Formally, we must show that for any fixed plan s_i' not t-equivalent to s_i and any rationalizable belief of $\tau^{s_i, t}$, the plan $\bar{s}_i$ gives a better expected payoff. Given a rationalizable belief on opponents' actions, player i has a uniform belief on the other players' actions as long as he follows s_i. Let $\hat{h}$ be a random variable equal to the shortest realized history at which s_i' differs from s_i before period t, or ∞ if they do not differ by period t. Note that the uniform belief on others' actions implies that $\hat{h} \neq \infty$ with positive probability. We show that conditional on any non-infinite value of $\hat{h}$, $\bar{s}_i$ strictly outperforms s_i' on average. In fact, this is weakly true date by date and strictly true at the first deviation, as shown by the following observations:

- At dates $1, \ldots, |\hat{h}|$, the plans are identical.
- At date $|\hat{h}| + 1$, the average payoff $f(\hat{h}, a_i)$ is strictly optimized by $\bar{s}_i(\hat{h})$.
- At dates $|\hat{h}| + 2, \ldots, t$, along the path observed by a player following s_i', the other players are known to repeat their date-$|\hat{h}| + 1$ move at dates $|\hat{h}| + 2, \ldots, t$. So at these dates, the plan s_i' cannot do better than to optimize with respect to the history truncated at length $|\hat{h}| + 1$. The plan $\bar{s}_i$ optimizes the expected stage-game payoffs with respect to a longer history, under which opposing moves are identical through date $|\hat{h}| + 1$. Since he is, therefore, solving a less constrained optimization problem, he must perform better than s_i' at each date $|\hat{h}| + 2, \ldots, t$.
- At dates $t + 1, \ldots$, under plan $\bar{s}_i$, player i now has complete information about his payoff and optimizes perfectly, so s_i' cannot do better.

If $\hat{h} = \infty$, again $\bar{s}_i$ cannot be outperformed because he optimizes based on complete information after t, and $\bar{s}_i$ and s_i' prescribe the same behavior before t.

Finally, since there are only finitely many histories and types in the construction, all payoffs are bounded and so can be normalized to lie in $[0, 1]$. □

Submitted 2014-6-15. Final version accepted 2015-1-22. Available online 2015-1-23.