COMMENT ON “SMITH (1995): PERFECT FINITE HORIZON FOLK THEOREM”

GHISLAIN-HERMAN DMEZE-JOUATS A \textsuperscript{a,2} AND ANDREA WILSON B,3

Smith (1995) proved a perfect folk theorem for finitely-repeated stage games with \textit{recursively distinct} Nash payoffs, without assuming a non-equivalent utilities (\textsc{neu}) condition. While his theorem is correct, the constructive proof contained a small gap, using strategies only guaranteed to form a \textsc{spne} under \textsc{neu}. Here, we illustrate the gap with a counterexample, and resolve it with a small adjustment to his strategies.

\textbf{Keywords:} repeated games, folk theorem, nonequivalent utilities.

1. \textsc{Introduction}

Benoit and Krishna (1985) proved a perfect finite-horizon folk theorem under a full dimensionality condition, assuming at least two distinct Nash equilibrium (\textsc{ne}) payoffs for each player. Smith (1995) extended their result by relaxing both assumptions: he showed that it is enough that the stage game have \textit{recursively distinct} \textsc{ne} payoffs, and allowed for players with affinely equivalent utilities.

His proof used a five-phase strategy profile, in which early deviations by player \textit{i} were punished using his \textit{effective minmax profile}: determine the highest minmax payoff among players equivalent to \textit{i}, and play the corresponding action profile. The construction only included punishments for non-equivalent players who deviated during \textit{i}’s minmax phase. Smith then referenced his earlier working paper Smith (1994) for proof that the proposed strategies constituted a subgame perfect Nash equilibrium (\textsc{spne}). But that paper ruled out equivalent players; in this case, \textit{i}’s effective minmax reduces to his standard minmax, so that the player being minmaxed gains nothing by deviating.

With affinely equivalent players, \textit{i}’s effective minmax profile may have the property that player \textit{i} himself is not playing a myopic best response, in which case punishments must be added to deter deviations by player \textit{i} (along with his affine twins) during his own minmax phase. We illustrate this issue with a counterexample, propose a small modification to the punishment phases, and show that the adjusted strategy profile constitutes a \textsc{spne}.

2. Smith’s Folk Theorem

Let $G = (A_i, \pi_i; i = 1, 2, \ldots, n)$ be a finite normal form \textit{n}-player game, where $A_i$ is player \textit{i}’s set of mixed strategies over a finite action set, $A \equiv \times_{i=1}^n A_i$.

\footnotetext[1]{The first author found the error and suggested a fix; the second became a co-author during the review process after suggesting a simpler resolution.}

\footnotetext[2]{I am grateful to Christoph Kuzmics, Frank Riedel, Tim Hellmann, and Olivier Gossner for helpful comments.}

\footnotetext[3]{I acknowledge helpful input from my wonderful dog, Mavi.}

\footnotetext[4]{Center for Mathematical Economics, Bielefeld University; demeze_jouatsa@uni-bielefeld.de.}

\footnotetext[5]{Georgetown University, Economics Department; aw1020@georgetown.edu.}
and \( \pi_i : A \rightarrow \mathbb{R} \) is \( i \)'s utility function. We assume that players have access to a public randomization device. Let \( \mathcal{I} = \{1, 2, \ldots, n\} \) be the set of players, and for any player \( i \), let \( \mathcal{I}(i) \) be the set players affinely equivalent to \( i \); normalize payoffs so that \( \pi_i(\cdot) = \pi_j(\cdot) \) for all \( j \in \mathcal{I}(i) \). Player \( i \)'s effective minmax payoff is \( \min_{a_i \in A} \max_{j \in \mathcal{I}(i)} \max_{a_j \in A_j} u_i(a_i, a_{-j}) \), or equivalently, the highest minmax payoff among players \( j \in \mathcal{I}(i) \).\(^1\) Normalize every effective minmax payoff to zero. Let \( F^* \) be the feasible and (strictly) individually rational payoff set, i.e. the set of all feasible payoff vectors \( w \) with \( w_i > 0 \) \( \forall i \).

Given a subset of players \( \mathcal{J} = \{j_1, \ldots, j_m\} \) and a corresponding (possibly mixed) action profile \( a_{\mathcal{J}} = (a_{j_1}, a_{j_2}, \ldots, a_{j_m}) \), let \( G(a_{\mathcal{J}}) \) be the induced \( (n - m) \)-player game for players \( \mathcal{I} \setminus \mathcal{J} \) obtained from \( G \) by fixing the actions of players in \( \mathcal{J} \) at \( a_{\mathcal{J}} \). Define a Nash decomposition of \( G \) as an increasing sequence of \( h \geq 1 \) nonempty subsets of players from \( \mathcal{I} \), namely \( \emptyset = \mathcal{J}_0 \subset \mathcal{J}_1 \subset \cdots \subset \mathcal{J}_h \subseteq \mathcal{I} \) so that for \( g = 1, 2, \ldots, h \), action profiles \( e_{\mathcal{J}_{g-1}}, f_{\mathcal{J}_{g-1}} \) exist with corresponding Nash payoff vectors \( y(e_{\mathcal{J}_{g-1}}) \) of \( G(e_{\mathcal{J}_{g-1}}) \) and \( y(f_{\mathcal{J}_{g-1}}) \) of \( G(f_{\mathcal{J}_{g-1}}) \) different exactly for players in \( \mathcal{J}_g \setminus \mathcal{J}_{g-1} \); and for any \( i \in \mathcal{J}_g \setminus \mathcal{J}_{g-1} \), let \( z^{g,i} \) be \( i \)'s least-preferred action profile among those yielding payoff vectors \( y(e_{\mathcal{J}_{g-1}}) \) and \( y(f_{\mathcal{J}_{g-1}}) \). The game has recursively distinct Nash payoffs if there is a Nash decomposition with \( \mathcal{J}_h = \mathcal{I} \).

Smith’s main result is as follows, with \( \{G(\delta, T) \mid T \text{-fold } \delta\text{-discounted repetition of } G\} \):

**Theorem 1 (Smith):** Suppose that the stage game \( G \) has recursively distinct Nash payoffs. Then for the finitely-repeated game \( G(\delta, T) \), \( \forall u \in F^* \) and \( \forall \varepsilon > 0 \), \( \exists T_0 < \infty \) and \( \delta_0 < 1 \) so that \( T \geq T_0 \) and \( \delta \in [\delta_0, 1] \Rightarrow \exists \) a SPNE payoff vector \( v \) with \( \|v - u\| < \varepsilon \).

**Gap in Original Proof.**

Smith’s proof was constructive (see full strategies below, along with the required adjustment). In it, early deviations by player \( i \) were punished via “Phase 3”:

**Phase 3:** Play \( i \)'s effective minmax profile. If \( j \notin \mathcal{I}(i) \) deviates early, start Phase 4.

This opens the door to a profitable one-shot deviation — hence the given strategies may not constitute a SPNE — as illustrated by a counterexample. Consider the following 3-player stage game \( G \), in which P1 chooses rows (\( T \) or \( B \)), P2 chooses columns (\( \ell \) or \( r \)), and P3 chooses matrices (\( L \) or \( R \)):

\[
\begin{array}{ccc|cc}
L & & & & R \\
\ell & | & r & & \ell & | & r \\
T & -1, -1, 0 & 1, 1, 0 & & T & 2, 2, 2 & 3, 3, 3 \\
B & -1, -1, 0 & 0, 0, 0 & & B & 2, 2, 1 & 2, 2, 2 \\
\end{array}
\]

Players 1 and 2 earn the same payoff at every profile, and thus are affinely equivalent. Player 1’s minmax payoff is \(-1\) (achieved if he best-responds to \( (\ell, L) \)),

\(^1\)See footnote 5 in Smith (1994) for this equivalent formulation of Wen (1994)’s definition.
player 2's minmax payoff is 0 (achieved if he best-responds to \((B, L)\)), and so they share an effective minmax payoff of 0, via the effective minmax profile \(\tilde{w}^1 = \tilde{w}^2 \equiv (B, r, L)\).

In Smith’s construction, player 1’s punishment phase specifies playing \(\tilde{w}^1\) for some number of periods, during which deviations by players 1 and 2 are ignored. But observe that P1 himself is not myopically best-responding at \(\tilde{w}^1\), and so he has a profitable one-shot deviation: play \(T\) instead of \(B\). This raises his current-period payoff from 0 to 1, with no future consequences.  

This issue is easily resolved with a two part adjustment to Smith’s Phase 3 (after an early deviation by player \(i\)): First, instead of playing \(i\)’s effective minmax profile, play the solution \(w^i\) to \(i\)’s effective minmax problem, namely, a profile \(w^i\) that minimizes \(\max_{j \in I(i)} \max_{a_j \in A_j} u_i(a_j, w^i_j)\). (In words, choose the profile \(w^i\) that minimizes the best that any affine twin of \(i\) gets by best-replying to \(w^i\); in the counterexample, \(w^1 = w^2 = (B, \ell, L)\)). Second, deter Phase 3 deviations by players in \(I(i)\) by threatening to restart Phase 3. This deterrent works because profile \(w^i\) has the property that the best any player in \(I(i)\) can earn by deviating is his effective minmax payoff 0.

3. CORRECTED PROOF

We now provide Smith (1995)’s full strategies — with Phase 3 modified as above — and prove that the adjusted strategies constitute a SPNE. Following Smith, choose a target payoff vector \(u^* \in F^*\). Fix a Nash decomposition into player subsets \(J_g\) (\(g = 1, 2, \ldots, h\)), along with the corresponding action profiles \(e_{J_g-1}\) and \(f_{J_g-1}\), and corresponding distinct (for players \(i \in J_g\backslash J_g-1\)) Nash payoff vectors \(y(e_{J_g-1})\) of \(G(e_{J_g-1})\) and \(y(f_{J_g-1})\) of \(G(f_{J_g-1})\). Define \(c_g \equiv \min_{i \in J_g\backslash J_g-1} \|y(e_{J_g-1})_i - y(f_{J_g-1})_i\|\). Let \(y^g\) denote alternating between the action profile yielding \(y(e_{J_g-1})\) (in even periods) and \(y(f_{J_g-1})\) (in odd periods).

We now construct a 5-phase strategy profile. The phase length variables — namely \(q\) (Phase 3), \(r\) (Phase 4), and \(t_g(q + r)\) (\(g = 1, 2, \ldots, h\), Phases 2 and 5) will be chosen at the end of the construction, along with the reward vectors \(x_j\) (\(\forall j \in I\)) used in Phase 4. Early\(^3\) (late) deviations are those occurring up to (after) period \(T - t_h(q + r) - (q + r)\).

**Strategy Profiles.**

1. (Main Path) Play (possibly via public randomization) a profile \(a\) yielding the target payoff vector, \(u^*\), until period \(T - t_h(q + r)\). After an early

---

\(^2\)The first author’s original paper (Demeze-Jouatsa (2018)) noted further that in this game, Smith’s strategies may not even yield a NE: If the target payoff vector holds P1’s payoff close to his effective minmax, 0, then P1 will actually have an incentive to trigger his minmax phase — where he’s able to earn 1 — as often as possible.

\(^3\)So a deviation is “early” if there is still time to run Phases 3 and 4 before period \(T - t_h(q + r) + 1\), when Phase 2 begins.
deviation by $i$, go to Phase 3; after a late deviation by $i \in J_{g'}$, go to Phase 5.

2. (Good Recursive Nash) For $g = h, \ldots , 1$: Play $y^g$ in periods $T - t_g(q + r) + 1, \ldots , T - t_{g-1}(q + r)$. After a deviation by $i \in J_{g'}$ with $g' < g$, start Phase 5. (On-path, this phase runs during the final $t_h(q + r)$ periods).

3. (Adjusted Minmax Phase for $i$): Play $w^i$ for $q$ periods, where $w^i$ solves $i$’s effective minmax problem (rather than playing $i$’s effective minmax profile, as in Smith).

- If any $j \not\in I(i)$ deviates early, start Phase 4; if any $j \in J_{g'}$ deviates late, start Phase 5 with $i \leftarrow j$.

- [Addition to Smith’s construction] If any $j \in I(i)$ deviates early, set $i \leftarrow j$ and restart Phase 3.

Then set $j \leftarrow i$ and start Phase 4.

4. (Reward Phase) Play $x^j$ for $r$ periods. If any $i$ deviates early, restart Phase 3; if any $i \in J_{g'}$ deviates late, start Phase 5. Then return to Phase 1.

5. (Bad Recursive Nash) Play $z_{g',i}$ until period $T - t_{g'-1}(q + r)$. (If $j \in J_{g''}$ deviates, where $g'' < g'$, set $g' \leftarrow g''$ and $i \leftarrow j$ and restart Phase 5.) Then go to Phase 2.

So along the equilibrium path, the sequence of action profiles is

\[
\underbrace{a, \ldots , a}_{T - t_h(q+r) \text{ periods}} ; \underbrace{y^h, \ldots , y^h}_{s_h(q+r) \text{ periods}} ; \overbrace{y^{h-1}, \ldots , y^{h-1}}^{s_{h-1}(q+r) \text{ periods}} ; \underbrace{y^1, \ldots , y^1}_{s_1(q+r) \text{ periods}}
\]

Since we next choose phase lengths such that $t_h(q + r)$ doesn’t depend on $T$, payoffs converge to $u^*$ for $T$ sufficiently large.

**Phase Lengths and SPNE Verification**

Let $\rho$ be the largest gap between best and worst payoffs across all players in $G$. For Phase 4, let $x^1, x^2, \ldots , x^n$ be feasible payoff vectors such that $x^i \gg 0$ $\forall i \in I$, $x^i < x_j^i \forall j \notin I(i)$, $x^i = x^j \forall j \in I(i)$, and $x^i < u^*_i \forall i \in I$. (Such vectors exist following Abreu et al (1994)).

Phase lengths are as follows:

- choose $q$ (length of Phase 3) to deter one-shot deviations, namely so that for all players $i$,

\[
(3.1) \quad \rho < q \cdot x^i_j
\]

- choose $r$ (length of Phase 4) to deter deviations by players $j \notin I(i)$ during Phase 3: namely such that for all $i$ and $j \notin I(i)$,

\[
(3.2) \quad \rho + \max \{ 0, (q - 1) \cdot (u^*_j - \pi_j(w^i) ) \} < r(x^i_j - x^j_j)
\]
• for the final recursive NE phase, the lengths are determined as follows:
  For any number $k$, let $\psi_{g}(k)$ be the least even number above $2k\rho/c_{g}$, so
  that that a player $i \in J_{g}$ is willing to play $k$ periods of any action followed
  by $\psi_{g}(k)$ periods of $y^{g}$, if deviations switch each $y^{g}$ to $z^{g,i}$. Recursively
  define

  \[
  s_{h}(m) = \psi_{h}(m) \text{ and } (\forall g = 1, 2, \ldots, h-1) \ s_{g}(m) = \psi_{g}(m+s_{g+1}(m)+\cdots+s_{h}(m))
  \]

  Then set $t_{0}(m) = 0$ and $t_{g}(m) = s_{1}(m) + \cdots + s_{g}(m)$, for $g = 1, 2, \ldots, h$.

  To prove that the strategies form a SPNE, it suffices to prove that there are no
  profitable one-shot deviations. We show that deviations are strictly unprofitable
  at $\delta = 1$, and thus remain unprofitable for $\delta$ sufficiently large.

  • Late deviations. A one-shot deviation by player $i \in J_{g'}$ takes him immediately
    to Phase 5, where they play $z^{g',i}$ until period $T - t_{g'-1}(q + r)$, then
    resume following Phase 2. So he gains at most $\rho$ in each period between
    the deviation and time $T - t_{g'}(q + r)$ (for a late deviation, there are at most
    $q + r + s_{h}(q + r) + s_{h-1}(q + r) + \cdots + s_{g'+1}(q + r)$ such periods), but then
    loses at least $c_{g}/2$ in each of the $s_{g'}(q + r)$ periods between $T - t_{g'}(q + r) + 1$
    and $T - t_{g'-1}(q + r)$ (during which they switch from $y^{g'}$ to $z^{g',i}$). By (3.3),
    the loss strictly exceeds the gain. (This analysis applies to late deviations
    by any player in Phases 1, 3, 4; to late Phase 2 deviations by players in $J_{g'}$
    from $y^{g'}$ (with $g > g'$); and to late Phase 5 deviations by players in $J_{g''}$
    from $z^{g',i}$ (with $g' > g''$). Remaining late deviations, by those already (by
    construction) playing a myopic best response, are ignored).

  • Early deviations in Phases 1 and 4. If $i$ deviates, he gains at most $\rho$ this
    period, then play moves immediately to Phase 3 (followed by Phase 4 with
    $x^{i}$). Since $x^{i}$ is weakly worse for player $i$ than any other Phase 4 vector
    $x^{i}$, and strictly worse than the Phase 1 vector $u^{*}$, the cost is at least $q \cdot x^{i}
    (i$ loses at least $x^{i}$ during each of the $q$ minmax periods). By (3.1),
    the deviation is unprofitable.

  • Early deviations by non twins during Phase 3 (minmaxing $i$). Player $j \in I(i)$
    gains at most $\rho$ in the current period, and then moves immediately to
    Phase 4, where he gets $x^{j}$ rather than the $x^{j}$ he would have gotten
    without the deviation. Then returns to Phase 1, so can replace at most
    $(q - 1)$ periods of minmaxing $i$ with payoff $u^{j}_{j}$. By (3.2), the deviation
    is unprofitable.

  • Early deviations by twins during Phase 3 (minmaxing $i$). A one-shot devi-
    ation by $j \in I(i)$ raises his payoff in the current period from $\pi_{j}(w^{i})$ to
    at best his effective minmax, zero. But this restarts Phase 3, adding at least
    one extra minmax period (at the expense of a future Phase 1 period), for
    a cost of at least $u^{j}_{j} - \pi_{j}(w^{i})$. Since $u^{j}_{j} > 0$, the deviation is unprofitable.
REFERENCES


