Identification of Payoffs in Repeated Games*

Byung Soo Lee†
University of Toronto

Colin Stewart‡
University of Toronto

Friday 13th May, 2016

Abstract

In one-shot games, an analyst who knows the best response correspondence can only make limited inferences about the players' payoffs. In repeated games with full monitoring, this is not true: we show that, under a weak condition, if the game is repeated sufficiently many times and players are sufficiently patient, the best response correspondence completely determines the payoffs (up to positive affine transformations).

1 INTRODUCTION

How much can one infer about players' payoffs in a game based only on their best response correspondences? In static games, such inferences are quite limited; while best responses convey some information about a player's preferences over her own actions for any given profile of the other players' actions, they say nothing about that player's preferences as the others' actions vary. Among other things, this makes welfare comparisons essentially impossible: one can show that for any profile of best response correspondences and any

*We are grateful to the editor, Françoise Forges, and two anonymous referees for helpful comments, and to Young Wu for excellent research assistance. Stewart thanks SSHRC for financial support.

†email: byungsoolee@gmail.com

‡email: colinbstewart@gmail.com

	L	R
T	a	b
B	c	d

	L	R
T	1	θ
B	0	$\theta + \frac{p^*}{1-p^*}$

Figure 1: A simple 2×2 game. Payoffs are for the row player. The matrix on the right illustrates the extent to which payoffs can be identified if the game is not repeated.

action profile $\mathbf{a}$ in a finite game, there exist payoffs according to which $\mathbf{a}$ is Pareto efficient, and payoffs according to which $\mathbf{a}$ is Pareto dominated, both of which lead to the given best response correspondences.

In repeated games with full monitoring, one can potentially infer much more. To the extent that other players' future actions depend on one's own current action, best responses convey information about preferences over others' actions. We show that this can be enough to fully identify payoffs (up to positive affine transformations). More precisely, as long as no player has an action ensuring that—regardless of others' actions—she obtains her highest possible payoff, the best response correspondences uniquely determine the payoffs when the game is repeated sufficiently many times and players are sufficiently patient.[1]

To illustrate, consider the 2×2 stage game depicted in Figure 1. First suppose this game is played once, and the best response correspondence for the row player is such that T is a best response if and only if the probability p that the column player assigns to L is at least $p^* \in (0, 1)$, while B is best response if and only $p \leq p^*$. What can we infer about the payoffs? First, $a > c$ and $d > b$. Without loss of generality, through an appropriate positive affine transformation, we may normalize a and c to be 1 and 0, respectively. Second, given this normalization, the row player's indifference between T and B when the column player chooses $p = p^*$ implies that $d - b = p^*/(1 - p^*)$. These conditions determine payoffs up to a constant parameter capturing the

[1] In the terminology of Myerson (2013), best-response equivalence implies full equivalence under these conditions.

row player's preferences across the two columns, and capture all that can be inferred from best responses. In particular, there is a continuum of distinct games having this best response correspondence.

Now suppose the game is played twice without discounting. Consider the row player's best responses to strategies that play L in the first period, followed by L in the second period if the row player played T in the first period, and a mixture assigning probability p to L and $1-p$ to R otherwise. Suppose we observe that the row player is indifferent between T and B in the first period when the column player uses this strategy with $p = p^{**}$. If $p^{**} > p^*$, then T is the best response for the row player in the second period regardless of her first-period action. Hence the indifference condition is

$$1 + 1 = 0 + p^{**} + (1 - p^{**})\theta,$$

from which we obtain $\theta = (2-p^{**})/(1-p^{**})$. Thus we can pin down the exact payoffs. This approach succeeds whenever $\theta > (2 - p^*)/(1 - p^*)$, ensuring that p^{**} is indeed greater than p^*.

By varying the column player's strategy and checking for indifferences for the row player between her first-period actions, one can identify θ in this way regardless of its value. More generally, however, if there is no strategy for the column player that makes the row player indifferent in the one-shot game, then more periods may be needed. For example, if in the game depicted in Figure 1 we have $a = 1$, $b = 3/2$, $c = 0$, and $d = 1/2$, then three periods are needed; with only two periods, varying the column player's action in the second period does not provide a strong enough incentive for the row player ever to prefer B in the first period. In general, although the number of repetitions needed depends on the payoffs, one can see from the best responses whether the payoffs can be identified.

This example is relatively simple, in part because the row player's payoffs can be identified in the static game up to the addition of a constant to each outcome in one column. In general, this may not be possible, as Figure 2

	L	C	R
T	2	2	2
M	0	3	0
B	3	0	0

	L	C	R
T	2	2	2
M	0	3	0
B	4	-2	-2

Figure 2: Two games with the same best response correspondence for the row player (when the game is not repeated).

indicates: the row player's set of best responses to any mixed strategy of the column player is identical in the two games depicted in the figure, but neither game can be obtained from the other by adding constants to columns.[2]

A number of papers have examined the testable restrictions of equilibrium notions in certain classes of games with various assumptions about what is observable to the analyst (Bossert and Sprumont, 2013; Chambers, Echenique, and Shmaya, 2010; Ledyard, 1986; Ray and Zhou, 2001; Ray and Snyder, 2013; Sprumont, 2000). We depart from this line of work in several respects. First, we take the game form as fixed and focus on identification rather than testable restrictions. Second, we do not assume that only equilibrium play is observable.[3] Experimental evidence suggests that subjects are often rational in the sense that they maximize expected utility with respect to some belief, but do not form correct beliefs about others' strategies (see, e.g., Costa-Gomes and Crawford (2006); Kneeland (2015)). In this case, although players may not play Nash equilibrium, best responses can be observed if beliefs are elicited (as in Nyarko and Schotter (2002)) or determined by experimental design (as in Agranov, Potamites, Schotter, and Tergiman (2012)). Although the assumption that the analyst can observe the full best response correspondence is quite strong, as we discuss below, our results require only knowledge of best responses to a small class of strategies. We do, however, require that the analyst know the extensive form structure of the

[2] Morris and Ui (2004) discuss a similar example.

[3] Abito (2015) studies partial identification of payoffs in repeated games based on equilibrium play. Nishimura (2014) considers the testable implications of individual rationality in extensive games when other players may not be rational.

game (in particular, that payoffs are constant across repetitions of the stage game).

Our work can be viewed as a strategic analogue of the classical problem of identifying preferences based on choices from menus (see, e.g., Arrow (1959)). In the classical model, if the set of menus is rich enough, one-shot choices are sufficient to fully identify preferences. The strategic structure of our setting effectively limits the kinds of menus from which the agent can choose, in which case making future menus contingent on the agent's choice can help to recover more information about preferences.

2 SETUP

2.1 Stage Game

A stage game is a tuple (I, A, g), where $A \coloneqq \prod_{i \in I} A_i$ and $g = (g_i)_{i \in I}$ such that

(i) I denotes the set of players;

(ii) for all $i \in I$, A_i denotes player i's action set; and

(iii) for all $i \in I$, $g_i \colon A \to \mathbb{R}$ denotes player i's payoff function.

We assume that I and each A_i are finite and have at least two members.

For the remainder, we fix I and A so that each stage game is identified by its payoffs g. We also let $\mathbf{A}_i \coloneqq \Delta(A_i)$ denote the set of i's mixed actions and let $\mathbf{A} \coloneqq \prod_{i \in I} \mathbf{A}_i$ denote the set of all mixed action profiles. We frequently abuse notation by identifying mixed action profiles with the product measure generated by them. Furthermore, let $g_i \colon A \to \mathbb{R}$ be extended to $g_i \colon \mathbf{A} \to \mathbb{R}$ in the usual way.[4]

[4] Thus $g_i(\mathbf{a}) = \sum_{a \in A} g_i(a)\mathbf{a}(a)$ for all $\mathbf{a} \in \mathbf{A}$.

Let $BR_i^1(\mathbf{a}_{-i})$ denote the set of i's pure action best responses to the opponents' mixed action profile $\mathbf{a}_{-i} \in \mathbf{A}_{-i}$; that is,

$$BR_i^1(\mathbf{a}_{-i}) \coloneqq \{a_i \in A_i \mid g_i(a_i, \mathbf{a}_{-i}) = \max_{b_i \in A_i} g_i(b_i, \mathbf{a}_{-i})\}.$$

2.2 Repeated Game

Now suppose that the game is played T times, where $1 \leq T \leq \infty$. Let $H^{<T} \coloneqq \bigcup_{t=0}^{T-1} A^t$ denote the set of all partial histories, where A^0 is a singleton consisting only of the empty history $\emptyset$. For any $\tau \geq 1$, $t \leq \tau$, and $h \in A^\tau$, let h^t denote the projection of the t-coordinate so that $h = (h^t)_{t=1}^\tau$. Furthermore, let h_i^t denote the i-coordinate of h^t so that $h^t = (h_i^t)_{i \in I}$.

Player i's (pure) strategy set in this repeated game is $S_i \coloneqq A_i^{H^{<T}}$ (i.e., a strategy is a map $s_i \colon H^{<T} \to A_i$ that specifies the planned action at each partial history). Let $\mathbf{S}_i \coloneqq \Delta(S_i)$ denote the set of i's mixed strategies, and let $S \coloneqq \prod_{i \in I} S_i$ and $\mathbf{S} \coloneqq \prod_{i \in I} \mathbf{S}_i$.

Let $\omega \colon S \to A^T$ map each strategy profile to the outcome that it induces. Furthermore, for all $1 \leq t < T$, let ω_t denote the t-coordinate of ω, i.e., if $\omega(s) = (a^t)_{t=1}^T$, then $\omega_t(s) = a^t$.

Player i's payoffs are given by the map $g_i^T \colon S \to \mathbb{R}$ defined by

$$g_i^T(s) \coloneqq \sum_{t=1}^{T} \delta_i^{t-1} g_i(\omega_t(s)) \tag{1}$$

for all $s \in S$, where $\delta_i \in (0, 1]$ and $\delta_i < 1$ if $T = \infty$.[5]

We extend the maps $\omega \colon S \to A^T$ to $\omega \colon \mathbf{S} \to \Delta(A^T)$, $\omega_t \colon S \to A$ to $\omega_t \colon \mathbf{S} \to \mathbf{A}$, and $g_i^T \colon S \to \mathbb{R}$ to $g_i^T \colon \mathbf{S} \to \mathbb{R}$ in the usual way.

The normal form of the repeated game is (I, S, g^T). Let $BR_i^T(\mathbf{s}_{-i})$ denote the set of i's pure strategy best responses to the opponents' mixed strategy

[5] If $\delta_i = 1$, we adopt the convention that $\frac{1-\delta_i^t}{1-\delta_i} = t - 1$.

profile $\mathbf{s}_{-i} \in \mathbf{S}_{-i}$; that is,

$$BR_i^T(\mathbf{s}_{-i}) \coloneqq \{s_i \in S_i \mid g_i^T(s_i, \mathbf{s}_{-i}) = \max_{r_i \in S_i} g_i^T(r_i, \mathbf{s}_{-i})\}.$$

2.3 Interactive Payoffs

Definition 2.1. *An action $a_i \in A_i$ is an* always-best response *if $a_i \in BR_i^1(\mathbf{a}_{-i})$ for every $\mathbf{a}_{-i} \in \mathbf{A}_{-i}$.*

Definition 2.2. *A payoff function g is* interactively determined *for player i (or i-*interactive*) if there is no always-best response a_i^* such that*

$$\min_{\mathbf{a}_{-i}} g_i(a_i^*, \mathbf{a}_{-i}) = \max_{\mathbf{a}_{-i}} g_i(a_i^*, \mathbf{a}_{-i}).$$

To put it another way, g is i-interactive if i cannot unilaterally determine her own payoff while choosing optimally; her payoff is determined jointly by her own (optimal) action and the actions of the other players. This is a very weak condition: it is violated only if there is some action for i in the stage game that, regardless of the other players' actions, gives i her maximum payoff in the game.

Whether g is i-interactive can be confirmed based on the best response correspondence in the repeated game. For any given δ_i and T, if there does not exist an action a_i for i such that the strategy $s_i(h) \equiv a_i$ is a best response to *every* strategy profile s_{-i}, then g is i-interactive. Conversely, if g is i-interactive, then there exist pairs (δ_i, T) for which this last condition holds. However, even if g is i-interactive, if i has an always best response in the stage game, then there are values of δ_i and T for which this is not possible to verify based only on the best response correspondence.

Lemma 2.3. *A payoff function g is i-interactive if and only if*

$$\min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) \neq \max_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}).$$

Proof. Suppose that g is *not* i-interactive. Then there is some a_i^* such that

(i) for all $\mathbf{a}_{-i} \in \mathbf{A}_{-i}$, $g_i(a_i^*, \mathbf{a}_{-i}) = \max_{a_i} g_i(a_i, \mathbf{a}_{-i})$; and

(ii) $\min_{\mathbf{a}_{-i}} g_i(a_i^*, \mathbf{a}_{-i}) = \max_{\mathbf{a}_{-i}} g_i(a_i^*, \mathbf{a}_{-i})$.

It follows that $\min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) = \max_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i})$.

For the other direction, suppose that

$$\min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) = \max_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}).$$

If there is some always-best response a_i^*, then the result follows immediately from the above equality.

Suppose by way of contradiction that there is no always-best response. Then, for all (a_i^*, a_{-i}^*) such that $a_i^* \in BR_i^1(a_{-i}^*)$, there is some b_{-i}^* such that $a_i^* \notin BR_i^1(b_{-i}^*)$.[6] Thus

$$\begin{aligned}
\min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) &= \max_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) \\
&= g_i(a_i^*, a_{-i}^*) \\
&= g_i(b_i^*, b_{-i}^*) \\
&> g_i(a_i^*, b_{-i}^*).
\end{aligned}$$

It follows that against any mixed action profile $\mathbf{a}_{-i}^+ = (\mathbf{a}_k^+)_{k \neq i}$ made up of full-support mixed actions,

$$g_i(\cdot, \mathbf{a}_{-i}^+) < \min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) = \max_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}), \tag{2}$$

which implies that

$$\min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}) \leq \max_{a_i} g_i(a_i, \mathbf{a}_{-i}^+) < \min_{\mathbf{a}_{-i}} \max_{a_i} g_i(a_i, \mathbf{a}_{-i}),$$

[6] Otherwise, a_i^* would be a best response to every pure action profile, and hence also a best response to every mixed action profile (i.e., it would be an always-best response).

a contradiction. ☐

3 RESULTS

We begin by examining the case in which the discount factor δ_i is known, and show that if δ_i is large enough, the payoffs of player i can be identified with enough repetitions of the game as long as they are i-interactive. Proposition 3.5 shows that if δ_i is unknown, it too can be identified from the best response correspondence.

The main result (Proposition 3.3) follows from a key lemma (Lemma 3.1).

Lemma 3.1. *Fix any $a^*_{-i}, a_{-i} \in A_{-i}$ and $j \neq i$ such that $a^*_k \neq a_k$ if and only if $k = j$. Furthermore, suppose that $a_i, a^*_i, b^*_i, c^*_i \in A_i$ and $\mathbf{b}^*_{-i}, \mathbf{c}^*_{-i} \in \mathbf{A}_{-i}$ satisfy the following:*

(i) $a^*_i \in BR^1_i(a^*_{-i})$, $b^*_i \in BR^1_i(\mathbf{b}^*_{-i})$, *and* $c^*_i \in BR^1_i(\mathbf{c}^*_{-i})$*;*

(ii) $0 < g_i(b^*_i, \mathbf{b}^*_{-i}) - g_i(c^*_i, \mathbf{c}^*_{-i})$*; and*

(iii) $g_i(a^*_i, a^*_{-i}) - g_i(a_i, a^*_{-i}) \leq \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b^*_i, \mathbf{b}^*_{-i}) - g_i(c^*_i, \mathbf{c}^*_{-i}))$.

Then there exists a unique $p \in [0,1]$ such that

$$g_i(a^*_i, a^*_{-i}) - g_i(a_i, a^*_{-i}) = p\frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b^*_i, \mathbf{b}^*_{-i}) - g_i(c^*_i, \mathbf{c}^*_{-i})). \quad (3)$$

In particular, for all $\hat{p} \in [0,1]$,

$$\hat{p} \geq p \iff \exists s_i \in BR^T_i(\mathbf{s}^{\hat{p}}_{-i}) \text{ with } s_i(\emptyset) = a_i \quad (4)$$

$$\text{and} \quad \hat{p} \leq p \iff \exists r_i \in BR^T_i(\mathbf{s}^{\hat{p}}_{-i}) \text{ with } r_i(\emptyset) = a^*_i, \quad (5)$$

where $\mathbf{s}^{\hat{p}}_{-i}$ *is defined by*

$$\mathbf{s}^{\hat{p}}_k(h) \coloneqq \begin{cases} a^*_k & \textit{if } h = \emptyset \\ a^*_k & \textit{if } h \in A^1 \wedge k \neq j \\ \hat{p}a^*_j + (1-\hat{p})a_j & \textit{if } h \in A^1 \wedge k = j \\ \mathbf{b}^*_k & \textit{if } h^1 = (a_i, a^*_{-i}) \wedge h^2_{-i} = a^*_{-i} \\ \mathbf{c}^*_k & \textit{if } h^1 \neq (a_i, a^*_{-i}) \vee h^2_{-i} \neq a^*_{-i}. \end{cases} \tag{6}$$

Proof. Let p be the unique solution to the equation

$$g_i(a^*_i, a^*_{-i}) - g_i(a_i, a^*_{-i}) = p\frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b^*_i, \mathbf{b}^*_{-i}) - g_i(c^*_i, \mathbf{c}^*_{-i})).$$

For any $h \in A^1$, let $r_i(h) = s_i(h) \in BR^1_i(pa^*_{-i} + (1-p)a_{-i})$. For any h such that $h^1 = (a_i, a^*_{-i})$ and $h^2_{-i} = a^*_{-i}$, let $r_i(h) = s_i(h) = b^*_i$. For any h such that $h^1 \neq (a_i, a^*_{-i})$ or $h^2_{-i} \neq a^*_{-i}$, let $r_i(h) = s_i(h) = c^*_i$. At such partial histories, i can do no better than choosing the stage game best response since the action profiles planned under $\mathbf{s}^p_{-i}$ at all subsequent partial histories do not depend on that choice.

Now let $r_i(\emptyset) = a^*_i$ and $s_i(\emptyset) = a_i$ so that the expected utilities for r_i and s_i against $\mathbf{s}^p_{-i}$ are, respectively,

$$g^T_i(r_i, \mathbf{s}^p_{-i}) = g_i(a^*_i, a^*_{-i}) + \delta_i \max_{d_i} g_i(d_i, pa^*_{-i} + (1-p)a_{-i}) \\ + \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i} g_i(c^*_i, \mathbf{c}^*_{-i})$$

and

$$g^T_i(s_i, \mathbf{s}^p_{-i}) = g_i(a_i, a^*_{-i}) + \delta_i \max_{d_i} g_i(d_i, pa^*_{-i} + (1-p)a_{-i}) \\ + \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(pg_i(b^*_i, \mathbf{b}^*_{-i}) + (1-p)g_i(c^*_i, \mathbf{c}^*_{-i})).$$

Then the difference $g_i^T(r_i, \mathbf{s}_{-i}^p) - g_i^T(s_i, \mathbf{s}_{-i}^p)$ equals

$$g_i(a_i^*, a_{-i}^*) - g_i(a_i, a_{-i}^*) - p\frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b_i^*, \mathbf{b}_{-i}^*) - g_i(c_i^*, \mathbf{c}_{-i}^*)) = 0.$$

It follows that $r_i, s_i \in BR_i^T(\mathbf{s}_{-i}^p)$. The remainder is immediate from the strict monotonicity (in p) of the left-hand side in the equation above. □

Definition 3.2. *Given δ_i, a best response correspondence BR_i^T* reveals *g_i if, up to positive affine transformations, there is a unique payoff function for which BR_i^T is the best response correspondence.*

Proposition 3.3. *Let g be i-interactive. Given any $\delta_i \in (0,1]$ sufficiently close to 1 and T sufficiently large, BR_i^T reveals g_i.*

Proof. Suppose there exist some $a_i^0, \bar{a}_i^0, \bar{a}_{-i}^0$ such that $g_i(\bar{a}_i^0, \bar{a}_{-i}^0) - g_i(a_i^0, \bar{a}_{-i}^0) > 0$ and $\bar{a}_i^0 \in BR_{-i}^1(\bar{a}_{-i}^0)$; if there are no such $a_i^0, \bar{a}_i^0, \bar{a}_{-i}^0$, then i-interactivity ensures that there exist strategies for which the corresponding inequality holds in the twice-repeated game, in which case the following argument goes through with the twice-repeated game as the stage game to show that g_i^2 is revealed by BR_i^T for sufficiently large T and δ_i, and hence so is g_i.

Choose $\hat{m}_i$, $\hat{\mathbf{m}}_{-i}^*$, $\check{m}_i$, and $\check{\mathbf{m}}_{-i}^*$ so that

(i) $\hat{m}_i \in BR_i^1(\hat{\mathbf{m}}_{-i}^*)$ and $\check{m}_i^* \in BR_i^1(\check{\mathbf{m}}_{-i}^*)$;

(ii) $g_i(\hat{m}_i, \hat{\mathbf{m}}_{-i}^*) = \max_{\mathbf{d}_{-i}} \max_{d_i} g_i(d_i, \mathbf{d}_{-i}) = \hat{M}$; and

(iii) $g_i(\check{m}_i, \check{\mathbf{m}}_{-i}^*) = \min_{\mathbf{d}_{-i}} \max_{d_i} g_i(d_i, \mathbf{d}_{-i}) = \check{M}$.

Because g is i-interactive, we have $\hat{M} > \check{M}$.

For any $a_i, a_i^*, b_i^*, c_i^* \in A_i$ and $a_{-i}^*, \mathbf{b}_{-i}^*, \mathbf{c}_{-i}^* \in A_{-i} \cup \{\check{\mathbf{m}}_{-i}^*, \hat{\mathbf{m}}_{-i}^*\}$, there exist sufficiently large $\delta_i \in (0,1]$ and T such that if $0 < g_i(b_i^*, \mathbf{b}_{-i}^*) - g_i(c_i^*, \mathbf{c}_{-i}^*)$, then $g_i(a_i^*, a_{-i}^*) - g_i(a_i, a_{-i}^*) \leq \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b_i^*, \mathbf{b}_{-i}^*) - g_i(c_i^*, \mathbf{c}_{-i}^*))$.[7] Furthermore,

[7] This is immediate from $\lim_{\delta_i \uparrow 1} \frac{\delta_i^2}{1-\delta_i} = +\infty$ and $\lim_{T\to\infty} \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i} = \frac{\delta_i^2}{1-\delta_i}$.

there exist sufficiently large $\delta_i \in (0,1]$ and T such that this property holds for all $a_i, a_i^*, b_i^*, c_i^* \in A_i$ and $\mathbf{b}_{-i}^*, \mathbf{c}_{-i}^* \in A_{-i} \cup \{\check{\mathbf{m}}_{-i}^*, \hat{\mathbf{m}}_{-i}^*\}$.[8] Fix such δ_i and T.

From Lemma 3.1, we can see that BR_i^T reveals the cardinal utility differences $g_i(b_i^*, b_{-i}^*) - g_i(b_i, b_{-i}^*)$ for all b_i, b_i^*, b_{-i}^* such that $b_i^* \in BR_i^1(b_{-i}^*)$. In fact, we can see that each such difference can be expressed as a scalar multiple of the difference $\hat{M} - \check{M}$, where the scalar $p(\cdot) \in [0,1]$ depends on (b_i, b_{-i}^*):

$$g_i(b_i^*, b_{-i}^*) - g_i(b_i, b_{-i}^*) = p(b_i, b_{-i}^*) \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(\hat{M} - \check{M}). \tag{7}$$

If $g_i(b_i^*, b_{-i}^*) - \check{M} = 0$, (7) gives

$$\begin{aligned} g_i(b_i, b_{-i}^*) &= g_i(b_i^*, b_{-i}^*) - p(b_i, b_{-i}^*) \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(\hat{M} - \check{M}) \\ &= \check{M} - p(b_i, b_{-i}^*) \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(\hat{M} - \check{M}). \end{aligned} \tag{8}$$

Now suppose $g_i(b_i^*, b_{-i}^*) - \check{M} > 0$. From Lemma 3.1, we can see that BR_i^T reveals the cardinal utility differences $g_i(\bar{a}_i^0, \bar{a}_{-i}^0) - g_i(a_i^0, \bar{a}_{-i}^0)$. In fact, for all (b_i^*, b_{-i}^*) such that $b_i^* \in BR_i^1(b_{-i}^*)$ and $g_i(b_i^*, b_{-i}^*) - \check{M} > 0$, the difference can be expressed as a scalar multiple of the difference $g_i(b_i^*, b_{-i}^*) - \check{M}$, where the scalar $q(\cdot) \in (0,1]$ depends on b_{-i}^*; $q(\cdot)$ is strictly positive because $g_i(b_i^*, b_{-i}^*) - \check{M} > 0$ and $g_i(\bar{a}_i^0, \bar{a}_{-i}^0) - g_i(a_i^0, \bar{a}_{-i}^0) > 0$. Thus we have

$$g_i(\bar{a}_i^0, \bar{a}_{-i}^0) - g_i(a_i^0, \bar{a}_{-i}^0) = q(b_{-i}^*) \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b_i^*, b_{-i}^*) - \check{M}).$$

Substituting (7) yields

$$q(b_{-i}^*) \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b_i^*, b_{-i}^*) - \check{M}) = p(a_i^0, \bar{a}_{-i}^0) \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(\hat{M} - \check{M}).$$

[8] For each $a_i, a_i^*, b_i^*, c_i^*, \mathbf{b}_{-i}^*, \mathbf{c}_{-i}^*$, find the smallest δ_i and T that work and take the maximum of all such δ_i and T since there finitely many of them.

Solving leads to

$$g_i(b_i^*, b_{-i}^*) = \check{M} + \frac{p(a_i^0, \bar{a}_{-i}^0)}{q(b_{-i}^*)}(\hat{M} - \check{M}).$$

Since the game can be normalized by letting $\check{M} = 0$ and $\hat{M} - \check{M} = 1$, the cardinal utility of any (b_i, b_{-i}^*) such that $g_i(b_i^*, b_{-i}^*) - \check{M} > 0$ (where $b_i^* \in BR_i^1(b_{-i}^*)$) can be expressed as

$$\begin{aligned} g_i(b_i, b_{-i}^*) &= g_i(b_i^*, b_{-i}^*) - p(b_i, b_{-i}^*)\frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i} \\ &= \frac{p(a_i^0, \bar{a}_{-i}^0)}{q(b_{-i}^*)} - p(b_i, b_{-i}^*)\frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}. \end{aligned} \tag{9}$$

Similarly, if $g_i(b_i^*, b_{-i}^*) - \check{M} = 0$, then by (8),

$$g_i(b_i, b_{-i}^*) = -p(b_i, b_{-i}^*)\frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}. \tag{10}$$

Let $\tilde{g}_i$ be a payoff function with the same best response correspondence BR_i^T as g_i. We claim that $\tilde{g}_i$ is an affine transformation of g_i. Indeed, if $\tilde{g}_i$ is normalized in the same way as g_i (so that $\check{M} = 0$ and $\hat{M} - \check{M} = 1$), then by the same derivation as for g_i, $\tilde{g}_i$ must satisfy (9) and (10), and by (4) and (5), $p(\cdot)$ and $q(\cdot)$ are the same for $\tilde{g}_i$ as for g_i. It follows that g_i is revealed by BR_i^T. □

Proposition 3.3 indicates that if g is i-interactive, then i's payoffs can be pinned down completely based on her best responses in the repeated game (for sufficiently large δ_i and T). What if g is *not* i-interactive? In that case, for each T, BR_i^T reveals only which action profiles give i her maximum payoff in the game; identifying i's preferences among any other outcomes is impossible. In particular, repeating the game makes no difference for identification of payoffs.

Although Proposition 3.3 is stated in terms of the entire best response correspondence for player i, much less is needed to identify payoffs; all that is required are the best responses to one class of strategy profiles of the form described in (6). In these profiles (of i's opponents), there is a "target" period 1 action for player i. In period 2, some player $j \neq i$ randomizes between two actions with some probability $\hat{p}$ on a given action a_j^*, while the other players play as they did in period 1. Actions from period 3 onward are constant and take on one of two values, which might respectively be described as "good" and "bad" for player i in the sense that player i gets a higher payoff from best-responding to the former than she does from best-responding to the latter. If no player deviates from the given strategies, the profile that is played from period 3 onward is determined jointly by player i's period 1 action and the outcome of player j's period 2 randomization. The good action profile is played from period 3 onward if player i played the target action in period 1 and player j played a_j^* in period 2. Otherwise, the bad action profile is played. In other words, probability $\hat{p}$ of getting the good action profile is offered as an incentive to play the target action in period 1 instead of myopically best-responding.

Our final result shows that if δ_i is unknown, it too can be identified from the best response correspondence under the same conditions as the stage game payoffs.

Definition 3.4. *A best response correspondence BR_i^T* reveals δ_i *if there is unique discount factor such that BR_i^T is the best response correspondence for some payoff function.*

Proposition 3.5. *Let g be i-interactive. Given any $\delta_i \in (0,1]$ sufficiently close to 1 and T sufficiently large, BR_i^T reveals δ_i.*

Proof. Let $\hat{m}_i$, $\hat{\mathbf{m}}_{-i}^*$, $\check{m}_i$, $\check{\mathbf{m}}_{-i}^*$, $\hat{M}$, and $\check{M}$ be as in the proof of Proposition 3.3. Also as in the proof of Proposition 3.3, let δ_i and T satisfy

$$g_i(a_i^*, a_{-i}^*) - g_i(a_i, a_{-i}^*) \leq \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(g_i(b_i^*, \mathbf{b}_{-i}^*) - g_i(c_i^*, \mathbf{c}_{-i}^*)).$$

for all $a_i, a_i^*, b_i^*, c_i^* \in A_i$ and $\mathbf{b}_{-i}^*, \mathbf{c}_{-i}^* \in A_{-i} \cup \{\check{\mathbf{m}}_{-i}^*, \hat{\mathbf{m}}_{-i}^*\}$ such that $0 < g_i(b_i^*, \mathbf{b}_{-i}^*) - g_i(c_i^*, \mathbf{c}_{-i}^*)$.

Let a_i, a_i^*, and a_{-i}^* be such that $g_i(a_i^*, a_{-i}^*) - g_i(a_i, a_{-i}^*) > 0$ (if no such a_i, a_i^*, and a_{-i}^* exist, the following argument applies to the twice-repeated game, just as in the proof of Proposition 3.3). Consider (7) from the proof of Proposition 3.3. The value of $p(a_i, a_{-i}^*)$ depends on T, so let us indicate this dependence with a superscript as follows:

$$g_i(a_i^*, a_{-i}^*) - g_i(a_i, a_{-i}^*) = p^T(a_i, a_{-i}^*) \frac{\delta_i^2(1 - \delta_i^{T-2})}{1 - \delta_i}(\hat{M} - \check{M}). \tag{11}$$

In the proof, we used the fact that BR_i^T reveals $p^T(a_i, a_{-i}^*)$ when $g_i(a_i^*, a_{-i}^*) - g_i(a_i, a_{-i}^*) \leq \frac{\delta_i^2(1-\delta_i^{T-2})}{1-\delta_i}(\hat{M} - \check{M})$. It is clear from the inequality that if it holds for $T = T^*$, then it also holds for $T = T^* + 1$. Therefore, we have

$$\frac{p^{T^*+1}(a_i, a_{-i}^*)}{p^{T^*}(a_i, a_{-i}^*)} = \frac{1 - \delta_i^{T^*-2}}{1 - \delta_i^{T^*-1}}. \tag{12}$$

Furthermore, the constant ratio on the left-hand side is revealed by $BR_i^{T^*+1}$. Note that $p^{T^*+1}(a_i, a_{-i}^*)/p^{T^*}(a_i, a_{-i}^*)$ belongs to the unit interval because $0 < p^{T^*+1}(a_i, a_{-i}^*) < p^{T^*}(a_i, a_{-i}^*)$. Let $f(\delta_i) = (1 - \delta_i^{T^*-2})/(1 - \delta_i^{T^*-1})$ and $f(1) = \lim_{\delta_i \uparrow 1} f(\delta_i) = (T^* - 2)/(T^* - 1)$. Then $f\colon [0,1] \to [0,1]$ is a strictly decreasing function of δ_i. Hence the solution to (12) is unique. □

REFERENCES

J. M. Abito. How much can we identify from repeated games? Working paper, 2015.

M. Agranov, E. Potamites, A. Schotter, and C. Tergiman. Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior*, 75(2):449–463, 2012.

K. Arrow. Rational choice functions and orderings. *Economica*, 26(102): 121–127, 1959.

W. Bossert and Y. Sprumont. Every choice function is backwards-induction rationalizable. *Econometrica*, 81(6):2521–2534, 2013.

C. P. Chambers, F. Echenique, and E. Shmaya. General revealed preference theory. Caltech SS Working Paper 1332, 2010.

M. A. Costa-Gomes and V. P. Crawford. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5):1737–1768, 2006.

T. Kneeland. Identifying higher-order rationality. *Econometrica*, 83(5):2065–2079, 2015.

J. O. Ledyard. The scope of the hypothesis of Bayesian equilibrium. *Journal of Economic Theory*, 39(1):59–82, 1986.

S. Morris and T. Ui. Best response equivalence. *Games and Economic Behavior*, 49(2):260–287, 2004.

R. B. Myerson. *Game Theory.* Harvard University Press, 2013.

H. Nishimura. Individual rationality in collective choice. Working paper, 2014.

Y. Nyarko and A. Schotter. An experimental study of belief learning using elicited beliefs. *Econometrica*, 70(3):971–1005, 2002.

I. Ray and S. Snyder. Observable implications of Nash and subgame-perfect behavior in extensive games. *Journal of Mathematical Economics*, 49:471–477, 2013.

I. Ray and L. Zhou. Game theory via revealed preferences. *Games and Economic Behavior*, 37(2):415–424, 2001.

Y. Sprumont. On the testable implications of collective choice theories. *Journal of Economic Theory*, 93(2):205–232, 2000.