Inefficient Stage Nash is not Stable

Thomas W. L. Norman*

Magdalen College, Oxford

June 11, 2018

Abstract

It is shown that, in infinitely-repeated games between two arbitrarily patient players, strategy profiles with inefficient pure stage-Nash continuations are not strategically stable (Kohlberg and Mertens, *Econometrica* 1986, 54: 1003–1039). By contrast, a set of strategy profiles similar to the Prisoners' Dilemma's "perfect tit-for-tat" is "uniformly robust to perfect entrants" (Swinkels, *Journal of Economic Theory* 1992, 57: 333–342), and hence contains a strategically stable set. Moreover, this set satisfies iterated dominance and a version of forward induction, whilst its stable subset is admissible. *Journal of Economic Literature* Classification: C72.

Key Words: repeated games; strategic stability; evolutionary stability.

1 Introduction

Kohlberg and Mertens (1986) argue that Nash equilibrium—and indeed any point-valued solution concept—misses important aspects of rational game-theoretic decision-making that are captured by their set-valued notion of "strategic stability". Moreover, the defects of point-valued solution concepts are at their most acute in repeated games, where there are already indications of the refinement potential of strategic stability: Osborne (1990) shows that, in finitely repeated coordination games, among the set of pure outcome paths that consist of sequences of one-shot Nash equilibria, only those with nearly Pareto efficient payoffs are stable.[1]

*I thank All Souls College, Oxford. Email thomas.norman@magd.ox.ac.uk.

[1] Van Damme (1989) also offers examples of the sometimes dramatic effects of stability in finitely repeated games. Whilst I am not aware of applications of stability to infinitely repeated games, Aumann and Sorin (1989) select the optimal outcome of an infinitely repeated two-person game of common interests using a perturbation in which every strategy with finite memory is used with positive probability. Their solution concept is, however, single-valued. Demange (1992) also applies stability to a game of indefinite length.

	Cooperate	Defect
Cooperate	2 2	3 0
Defect	0 3	1 1

Figure 1: Prisoners' Dilemma

Game theorists have long been suspicious of repeated-game strategy profiles that are "unforgiving" in the sense that they can become locked into perpetual inefficiency. This paper shows that well-known unforgiving strategy profiles are strategically unstable with two arbitrarily patient players, whilst certain forgiving profiles are stable. In particular, sets of strategy profiles with inefficient pure stage-Nash continuations are not strategically stable. For instance, in the Prisoners' Dilemma of Figure 1, infinite repetition of the stage-Nash equilibrium is unstable, as is the famous "grim–trigger" equilibrium (Friedman 1971). By contrast, a set of strategy profiles similar to "perfect tit-for-tat" does contain a strategically stable set.

There is a simple intuition behind these results. In perpetual inefficient stage-Nash equilibrium, it is an alternative best reply for an arbitrarily patient player to experiment with efficient play, and if it is not reciprocated, revert to playing stage Nash; she can do no worse than under stage Nash if she reverts. Stable sets need not include alternative best replies, but there exists a perturbed game where the experimentation is reciprocated with positive probability—opponents make exactly the right mistakes—so that experimentation offers a profitable deviation. The same argument can be made of the stage-Nash continuation under grim–trigger strategies following a mistake, and in the perturbed games relevant for strategic stability, a mistake occurs with positive probability. By contrast, under perfect tit-for-tat, each player cooperates if and only if both players played the same action in the previous period. Hence, it has the property that it punishes defections (and failures to punish defections), but subsequently forgives them in the sense of returning to efficient play. This means that it punishes deviations that do not "play like" it (in a sense to be made precise) both on and off the equilibrium path, and hence it belongs to a set of strategy profiles that satisfies the evolutionary stability criterion of being "uniformly robust to perfect entrants" or URPE (Swinkels 1992), which in turn must contain a Kohlberg–Mertens stable set.

There are three noteworthy remarks to be made in formalizing this intuition. First, the strongest instability results are available under time-average payoffs (Section 3), although a weaker form of instability can be demonstrated in the discounting case (Section 4). Second, Kohlberg and Mertens (1986) define strategic stability for finite games, and hence I employ Al-Najjar's (1995) extension to games with infinitely many pure strategies. Among his desirable properties of stable sets, only admissibility extends straightforwardly to time-average payoffs, but iterated dominance and a version of forward induction do apply to URPE sets (Subsection 5.1). Third, uniform robustness to perfect entrants is also defined for finite games, and so I extend its definition to an infinite setting, along with Swinkels' (1992) result that certain such sets contain a stable set (Section 5). I begin, however, by reviewing the rich body of related literature.

Evolutionary stability In its traditional formulation, evolutionary stability has had limited success in selecting between the equilibria possible under the various Folk Theorems. Axelrod and Hamilton (1981) show that "always defect" is not an evolutionarily stable strategy (ESS) in the repeated Prisoners' Dilemma with time-average payoffs, since it is vulnerable to invasion by "tit-for-tat" (whereby a player cooperates in the first period and thereafter chooses the action her opponent took in the previous round), though this breaks down under discounting. Axelrod's (1981, 1984) celebrated evolutionary simulations of the repeated Prisoners' Dilemma found selection pressure in favor of tit-for-tat, but the outcome of such simulations is quite sensitive to the initial distribution of strategies upon which the selection process acts. On a theoretical level, Axelrod argues in favor of tit-for-tat as a "collectively stable strategy", but this concept does not imply evolutionary stability and gives little sharpening of the Nash Folk Theorem. Moreover, tit-for-tat is not a subgame perfect equilibrium strategy against itself, and thus is not even sustainable under the perfect Folk Theorems.

The usual formulation of evolutionary stability suffers from severe existence problems in infinitely repeated games (Boyd and Lorberbaum 1987, Farrell and Ware 1989, Kim 1994), whilst a switch to neutral stability gives little refinement of the predictions of the Folk Theorem. Boyd and Lorberbaum (1987) show that no pure strategy can be evolutionarily stable in the infinitely repeated Prisoners' Dilemma, whilst Farrell and Ware (1989) extend this to finite mixtures of pure strategies. Kim (1994) generalizes these results to any strategies, and also to Selten's (1983) extensive-form concept of direct ESS. But Sugden (1986) and Boyd (1989) show that ESS's do exist if players occasionally make mistakes. The existence problem for direct ESS is the possibility of mutation to strategies that differ from the existing ones only off the equilibrium path. Selten's notion of a limit ESS addresses this problem by perturbing the game—so that every information set is reached with positive probability—and finding the limit of a sequence of direct ESS's as the perturbations vanish. This gives a refinement of sequential equilibrium in symmetric extensive-form games (van Damme 1987). However, Kim proves that a Folk Theorem obtains for limit ESS's; the concept offers no sharpening of the predictions of subgame perfection in the infinitely repeated Prisoners' Dilemma.

A similar criticism can be levelled at the relaxation of evolutionary stability to neutral stability, even with time-average payoffs, where there exist neutrally stable strategies of the infinitely repeated Prisoners' Dilemma that are arbitrarily close to "always defect" for example (Fudenberg and Maskin 1990). Nonetheless, Fudenberg and Maskin (1990) and Binmore and Samuelson (1992) do find efficiency to be implied by modified versions of neutral stability. Binmore and Samuelson's incorporation of complexity costs into neutral stability destabilizes the off-the-equilibrium-path punishments required to prevent secret handshakes, and thus provides selection pressure in favor of efficiency. Fudenberg and Maskin (1990), meanwhile, demonstrate that when players employ finitely complex strategies and have time-average payoffs that are lexicographic in infinitesimally likely mistakes, neutral stability gives efficiency in the infinitely repeated Prisoners' Dilemma. The essential idea is that when players make mistakes, the worst possible history for an inefficient strategy profile will eventually occur—a defection under grim–trigger strategies, for instance. Such a profile is then vulnerable to invasion by a mutant that mimics the incumbent strategy except after this worst history, at which point it has nothing to lose from engaging in an evolutionary "secret handshake" (Robson 1990) that cooperates against itself but continues to defect against incumbents.

This idea is close to the logic in this paper. The Nash equilibrium refinement of (normal-form) perfection (Selten 1975) explicitly incorporates the mistakes that are implicit in Fudenberg and Maskin's model. However, perfection is inadequate to destabilize inefficiency, since it requires robustness only against a particular sequence of mistakes. Fudenberg and Maskin's symmetric evolutionary setting means that mutants are guaranteed to play themselves with positive probability, and outside of this setting the required mistakes are quite special, precisely reciprocating a player's secret handshake. By requiring robustness against all possible mistakes, however, we will see that strategic stability also selects against unforgiving strategies.

Strategic stability In a finite n-player game, a closed set M of Nash equilibria is stable (Kohlberg and Mertens 1986) if it is minimal with respect to the following property: for any $\varepsilon > 0$, there exists some β such that, for any completely mixed-strategy profile $(\phi_1, \ldots, \phi_n)$ and for any $(\beta_1, \ldots, \beta_n)$, $0 < \beta_i < \bar{\beta}$, the perturbed game where every strategy μ_i of player i is replaced by $\tilde{\mu}_i = (1-\beta_i)\mu_i + \beta_i\phi_i$ has an equilibrium ε-close to M. Intuitively, M is perfect under all possible perturbations. Such a set satisfies a list of desirable properties that includes existence, invariance, admissibility, iterated dominance and forward induction. For this reason, it is Kohlberg and Mertens' preferred notion of three stability concepts that they discuss. It does not, however, satisfy connectedness or backwards induction, whereas subsequent reformulated notions of stability do (Mertens 1989, 1991, Hillas 1990). Nonetheless, here I use the original Kohlberg–Mertens concept, because Swinkels' (1992) concept of a URPE set is closely tied to Kohlberg–Mertens stability alone.

Kohlberg and Mertens argue forcefully for the normal-form sufficiency view that a game-theoretic solution concept should depend only on a game's (reduced) normal form. Perhaps the most striking illustration of this view is the result that any proper equilibrium (Myerson 1978) of a normal-form game is a sequential equilibrium (Kreps and Wilson 1982b) of any extensive-form game with that normal form (van Damme 1984, Kohlberg and Mertens 1986).[2] Whilst normal-form sufficiency is not uncontroversial, strategic stability does provide a unified solution to numerous instances of troublesome behavior under point-valued extensive-form solution concepts: it implies the "intuitive criterion" of Cho and Kreps (1987), and coincides with a variation of universal divinity in an appealing class of signaling games (Cho and Sobel 1990); it rejects equilibria sustained by counterintuitive threats in a two-stage Prisoners' Dilemma–coordination game (Glazer and Weiss 1990); it adds a mixed Nash equilibrium to the paradoxical subgame perfect equilibrium of the centipede game (van Damme 1989, pp. 482–3); it implies the Kreps–Wilson (1982a) "plausible-beliefs" solution to the chain-store paradox (Govindan 1995); and it gives incumbent under-investment in the entry game with avoidable fixed costs (Bagwell and Ramey 1996), in line with the lack of empirical support for the classic Dixit (1980) prediction of over-investment.

A number of authors have drawn connections between strategic stability and evolutionary stability. Swinkels (1992) shows that a certain "shaped" set that is "uniformly robust

[2]Note that sequential equilibrium fails to satisfy invariance.

against equilibrium entrants" contains a stable set (variously defined), whilst the weaker concept of "uniform robustness against perfect entrants" suffices to contain a Kohlberg–Mertens stable set. A URPE set need only resist invasion by mutants playing a robust (normal-form perfect) best reply to the population strategy profile that their presence creates. An evolutionarily stable set (Thomas 1985), by contrast, must resist invasion by all mutants, and is hence a refinement of a URPE set. The latter's perfectness requirement on mutants limits the "drift" into alternative best replies that plagues evolutionary stability in infinitely repeated games. Swinkels (1993) shows that if a set meeting a certain topological condition is asymptotically stable in an evolutionary dynamic from a broad class, then it contains a hyperstable set of Nash equilibria (Kohlberg and Mertens 1986). Ritzberger and Weibull (1995) show that every face of strategies that is asymptotically stable in all evolutionary dynamics from a certain class contains a Kohlberg–Mertens stable set, whilst Demichelis and Ritzberger (2003) provide a condition that strengthens the inclusion to Mertens' (1989) reformulated stability.

2 Stability in Repeated Games

Consider a repeated game with two players $i \in I \equiv \{1, 2\}$, each with a finite pure action set A_i, from which she selects an action a_i in each of the periods $t \in \mathbb{N}$.[3] Letting A be the set $A_1 \times A_2$ of action profiles, payoffs in each stage game γ are given by a *payoff function* $u : A \to \mathbb{R}^2$. Let $\mathscr{G}$ denote the class of such repeated games Γ. Letting $\mathscr{C}(B)$ be the intersection of all convex sets containing the set B, a feasible payoff $v \in \mathscr{C}(u(A))$ is *strongly dominated* by another feasible payoff v' if $v'_i > v_i$ for both i. An action profile $a \in A$ is *weakly inefficient* if it yields the players a payoff that is strongly dominated by another feasible payoff. There is perfect monitoring of realized action profiles. A *(finite) history* $h^t \in \overline{H}^t$ is a list $(a^0, \ldots, a^{t-1})$ of the t action profiles played in periods 0 through $t-1$, with $\overline{H} \equiv \bigcup_{t=0}^{\infty} \overline{H}^t$ the set of all possible histories, indexed by $m = 1, 2, \ldots$. A *play* p is an infinite sequence $(a^0, a^1, \ldots)$ of action profiles, belonging to the space $\mathscr{P} \equiv A^\infty$. For every history $h^t \in \overline{H}$, a *cylinder with base on* h^t is the set $C(h^t) = \{p \in \mathscr{P} \mid p = (h^t, \ldots)\}$ of all realized plays whose t initial elements coincide with h^t. Let $\mathscr{F}^t$ be the σ-algebra on $\mathscr{P}$ whose elements are all finite unions of cylinders with base on $\overline{H}^t$. We then have a filtration

$$\mathscr{F}^0 \subset \cdots \subset \mathscr{F}^t \subset \cdots \subset \mathscr{F},$$

[3] Extending the results to n-player games would require a public randomization device, which I do not include here.

where $\mathscr{F}^0$ is the trivial σ-algebra and $\mathscr{F}$ is the σ-algebra generated by the algebra of histories $\mathscr{F}_0 \equiv \bigcup_{t\in\mathbb{N}} \mathscr{F}^t$.

Player i's set of *pure strategies* is the set X_i of mappings $x_i : \overline{H} \to A_i$. A pure-strategy profile $x \equiv x_1 \times x_2$ recursively induces the play $a(x) \equiv (a^0(x), a^1(x), a^2(x), \ldots) \in \mathscr{P}$. Given a history $h^t \in \overline{H}$, the *continuation* game is the infinitely repeated game that begins in period t, following history h^t; $x|_{h^t}$ is then the pure-strategy profile induced by x in the normal form of the continuation game following h^t. Taking the discrete topology on stage-game action profiles, the space X of pure-strategy profiles is endowed with the product topology over histories, in which it is compact by Tychonoff's theorem. This is a Cantor space (see Aliprantis and Border, 2007, p. 98), metrizable by

$$D(x,y) = \sum_{h\in\overline{H}} \frac{1 - \mathbf{1}_h(x,y)}{3^m},$$

where $\mathbf{1}_h(x,y)$ is the indicator function taking the value 1 if $a(x|_h) = a(y|_h)$, $h \in \overline{H}$, and 0 otherwise. Note that any cylinder set $C(h)$, $h \in \overline{H}$, is induced by an open set of pure strategies in the product topology on X, which will be crucial for a number of the results to follow. If, for all $h \in \overline{H}$ and all but finitely many $t \in \mathbb{N}$, $a^t(x|_h) = a^t(y|_h)$, then x is said to *play like* y.

A *mixed strategy* μ_i then belongs to the space Φ_i of mixtures of pure strategies. A mixed-strategy profile $\mu \equiv \mu_1 \times \mu_2$ belongs to the space $\Phi \equiv \Phi_1 \times \Phi_2$, and induces a probability measure ξ_μ on the set $\mathscr{P}$ of plays. In particular, ξ_μ is defined inductively on the cylinder sets, with $\xi_\mu(C(h))$ giving the probability of the history h; $\xi_\mu(C(\emptyset)) = 1$ and $\xi_\mu(C(ha)) = \xi_\mu(C(h)) \times \mu|_h(a)$, $h \in \overline{H}$, $a \in A$. The Kolmogorov extension theorem then delivers a unique extension of ξ_μ from the $\mathscr{F}^t$'s to $\mathscr{F}$. The mixed-strategy space Φ is compact in the weak* topology, and metrizable with metric

$$d(\mu,\mu') = \sum_{h\in\overline{H}} \frac{1 - \mathbf{P}_h(\mu,\mu')}{3^m},$$

where $\mathbf{P}_h(\mu,\mu')$ is the probability that μ and μ' induce the same continuation play $a(\cdot|_h)$ following history h.[4] In this space, the pure strategy x_i is denoted by the Dirac measure δ_{x_i}.

Given a pure-strategy profile $x \in X$ and a period t, the induced pure action profile

[4]In the weak* topology, a sequence $\{\mu^k\}$ in Φ converges to μ if, for every continuous function $f : X \to \mathbb{R}$, the sequence of real numbers $\int_X f\,d\mu^k$ converges to $\int_X f\,d\mu$. The mixed strategy space Φ is compact in the weak* topology by the Banach–Alaoglu Theorem.

$a^t(x)$ yields a flow payoff of $u_i(a^t(x))$ to player i. I will be concerned first of all with arbitrarily patient players, in the sense that each player i aggregates her flow payoffs over time according to the *time-average payoff* $l_i(x) \equiv \limsup_{T\to\infty} \sum_{t=0}^{T} u_i(a^t(x))/T$.[5] However, I will also be interested in (non-arbitrarily) patient players, who maximize their *average discounted payoffs*,

$$f_i(x) = (1-\rho)\sum_{t=0}^{\infty} \rho^t u_i(a^t(x)),$$

where $\rho \in [0,1)$ will be "sufficiently large". The corresponding expected payoffs to player i under the mixed-strategy profile μ are given by

$$\overline{U}_i(\mu) = \int_X l_i(x)\,d\mu \quad \text{and} \quad U_i(\mu) = \int_X f_i(x)\,d\mu.$$

Given a repeated game $\Gamma \in \mathscr{G}$, the normal form of Γ is defined by the triple (I, X, l) or (I, X, f), where $l \equiv (l_1, l_2)$ and $f \equiv (f_1, f_2)$.

A measure η_i on X_i is *positive* if $\eta_i(E) \geq 0$ for every measurable set E. It is *strictly positive* if it is positive and $\eta_i(E) > 0$ for every open set E. Al-Najjar (1995) defines a set of measures $\mathscr{M}_i$ on X_i as an *(admissible) perturbation class for player i* if it satisfies the following conditions:

1. $\mathscr{M}_i$ is a convex set of positive measures of norm no greater than 1;

2. $\mathscr{M}_i$ contains 0 and at least one strictly positive measure; and

3. for every Borel set B, and any perturbation η_i in $\mathscr{M}_i$, the restriction $\eta_{i[B]}$ of η_i to B is also in $\mathscr{M}_i$.

Let $\mathring{\mathscr{M}}_i$ denote the subset of $\mathscr{M}_i$ consisting of strictly positive measures. If a mixed strategy μ_i is such that $\mu_i \geq \eta_i$ for some $\eta_i \in \mathring{\mathscr{M}}_i$ (and setwise inequality of measures), then μ_i is called *completely mixed*. Given $\mu \in \Phi$, let μ_{-i} be the strategy μ_j of player i's opponent, player $j \neq i$. Given $\eta \in \mathscr{M} \equiv \mathscr{M}_1 \times \mathscr{M}_2$ and a mixed-strategy profile $\mu \in \Phi$, let $\overline{BR}_i(\mu, \eta) \equiv \{\phi_i \in \Phi_i \mid \forall \mu_i' \in \Phi_i : \mu_i' \geq \eta_i \Rightarrow \overline{U}_i(\phi_i, \mu_{-i}) \geq \overline{U}_i(\mu_i', \mu_{-i})\}$ be Al-Najjar's *restricted best-reply correspondence* under time-average payoffs, with $\overline{BR}(\mu, \eta) \equiv \overline{BR}_1(\mu, \eta) \times \overline{BR}_2(\mu, \eta)$ and $\overline{BR}_i(\cdot, 0) \equiv \overline{BR}_i(\cdot)$. Let $BR_i(\mu, \eta) \equiv \{\phi_i \in \Phi_i \mid \forall \mu_i' \in \Phi_i : \mu_i' \geq \eta_i \Rightarrow U_i(\phi_i, \mu_{-i}) \geq U_i(\mu_i', \mu_{-i})\}$, $BR(\mu, \eta)$ and $BR_i(\cdot, 0)$ be the analogous objects under average discounted payoffs.

[5]Either the limit superior or inferior has the advantage of existence over the classical limit, and the advantage of integrability over a Banach limit. However, I use the limit superior for its upper semicontinuity, which is important for existence of a best reply under time-average payoffs.

Definition 1 *(i) Suppose that* $\bar{\eta} \in \mathscr{M}$ *and that* $Y \subset \Phi$ *is an open set containing the compact set of equilibria* M*. Then we say that* M *is* $(Y, \bar{\eta})$-pre-stable under time-average payoffs (resp., average discounted payoffs) *if for every* $\eta \in \mathscr{M}$ *with* $\eta \leq \bar{\eta}$ *the correspondence* $\overline{BR}(\cdot, \eta)$ *(resp.,* $BR(\cdot, \eta)$*) has a fixed point* μ *in* Y*.*
(ii) A compact set of equilibria M *is* pre-stable under time-average payoffs (resp., average discounted payoffs) *if for every open* Y *containing* M *there exists an* $\bar{\eta} \in \mathring{\mathscr{M}}$ *such that* M *is* $(Y, \bar{\eta})$*-pre-stable under time-average payoffs (resp., average discounted payoffs).*
(iii) A compact set of equilibria M *is* stable under time-average payoffs (resp., average discounted payoffs) *if it is a minimal pre-stable set under time-average payoffs (resp., average discounted payoffs).*

Letting Γ_η be the game obtained from Γ by adding the restriction η on strategies, we can see that the set of Nash equilibria of Γ_η is the set of fixed points of the correspondence $\overline{BR}(\cdot, \eta)$ (resp., $BR(\cdot, \eta)$).

Since f_i is continuous, U_i is weak* continuous and linear in Φ_i by Al-Najjar's Proposition 2.1. Hence, we have existence of a stable set under average discounted payoffs by his Proposition 3.2, and his formulations of admissibility, iterated dominance and forward induction by his Propositions 4.1–4.4. By contrast, l_i is not continuous when X is endowed with the product topology. For two strategy profiles may be close in the product topology even if they are quite different in the distant future; in particular, $x, y \in X$ could induce different action profiles for all $t \geq \tau$, yet still be close in the product topology for τ sufficiently large. This is inconsistent with continuity of l_i, under which players care only about outcomes occurring infinitely often; if $\mu \in \Phi$ plays like $\mu' \in \Phi$, then $\overline{U}_i(\mu) = \overline{U}_i(\mu')$, by shift invariance of the limit superior. Al-Najjar's Propositions 2.1, 2.2 and 3.2 thus fail under time-average payoffs, so that we have no general existence result to invoke.[6] Nor does Carbonell-Nicolau's (2011) result for discontinuous games apply, since the class $\mathscr{G}$ of repeated games fails his Condition (B) under time-average payoffs. However, $\overline{BR}_i(\mu, \eta)$ is at least nonempty by upper semicontinuity of $\overline{U}_i$ and the appropriate extension of the Weierstrass Extreme Value Theorem.[7] Moreover, I will later demonstrate the existence of a stable subset of a URPE set in a class of repeated games that includes the Prisoners' Dilemma. We will see in Subsection 5.1 that this URPE set satisfies iterated dominance and a version of forward induction, whilst its stable subset continues to be admissible.

[6]The linearity of $\overline{U}_i$ does carry over from Al-Najjar's Proposition 2.1 by Fubini's Theorem (since l_i is integrable), and I will have frequent use for this property.

[7]The Berge Maximum Theorem applies only in part; see Leininger (1984) for details—his Lemma but not his Theorem applies here. See also Ausubel and Deneckere (1993).

3 Unstable Sets with Arbitrary Patience

Let $\Delta(X)$ be the set of probability measures on the set X. Given a stage-Nash equilibrium $\pi^* \in \Delta(A)$ of γ, let *stage Nash* be the strategy profile $\mu^* \in \Phi$ of Γ that plays π^* following any history $h \in \overline{H}$.

Lemma 1 *For any completely-mixed strategy profile $\mathring{\mu} \in \Phi$, any history $h \in \overline{H}$ has strictly positive probability under $\mathring{\mu}$, i.e. $\xi_{\mathring{\mu}}(C(h)) > 0$.*

This follows immediately from $C(h)$ being induced by an open set in the product topology on X, and a completely mixed-strategy profile putting strictly positive probability on any open set.

Proposition 1 *If $\pi^* \in \Delta(A)$ is a weakly inefficient pure equilibrium of γ, and $M \subseteq \Phi$ is a set of strategy profiles of Γ that play like stage Nash μ^*, then M is not stable under time-average payoffs.*

Proof. Since π^* is weakly inefficient, let $\bar{p} \in \mathscr{P}$ be a play that yields the players a time-average payoff by which π^* is strongly dominated. Let $\bar{x}$ be the pure-strategy profile that plays $\bar{p}$ following any history. Consider a strategy profile $\phi \in \Phi$ such that, for all t:

- $\phi|_{h^t} = \delta_{\bar{x}}$ if $h^t = (a^0(\bar{x}), a^1(\bar{x}), \ldots, a^t(\bar{x}))$; and
- $\phi|_{h^t} = \mu^*|_{h^t}$ otherwise.

Now consider a strategy profile $\mathring{\phi}^\kappa \in \Phi$ such that, for all $h \in \overline{H}$, $\mathring{\phi}^\kappa|_h = (1-\kappa)\phi|_h + \kappa\mathring{a}$, where $\kappa \in (0,1)$ and $\mathring{a}$ is a full-support mixture over action profiles (i.e. $\mathring{a}$ puts strictly positive probability on each $a \in A$). Then I claim that there exists $\bar{\kappa}$ sufficiently small that $\mathring{\phi}^{\bar{\kappa}}$ is a completely mixed-strategy profile with $\overline{U}_i(\phi_i, \mathring{\phi}^{\bar{\kappa}}_{-i}) > \overline{U}_i(\mu_i, \mathring{\phi}^{\bar{\kappa}}_{-i})$ for both i and all $\mu \in M$. To see this, note first that if $\mathring{\phi}^{\bar{\kappa}}$ were not completely mixed, there would exist an open set $E \subset X$ such that $\mathring{\phi}^{\bar{\kappa}}(E) = 0$, and hence some history $h \in \overline{H}$ with zero probability under $\mathring{\phi}^{\bar{\kappa}}$ (i.e. $\xi_{\mathring{\phi}^{\bar{\kappa}}}(C(h)) = 0$), contradicting $\mathring{a}$'s full support. Second, clearly $\overline{U}_i(\phi) > \overline{U}_i(\mu_i, \phi_{-i})$ for all $\mu \in M$, and $\overline{U}_i(\cdot, \mathring{\phi}^{\bar{\kappa}}_{-i}) \to \overline{U}_i(\cdot, \phi_{-i})$ as $\bar{\kappa} \to 0$.

Next consider the perturbed game $\Gamma_{\beta\mathring{\phi}^{\bar{\kappa}}}$, $\beta \equiv (\beta_1, \beta_2)' \in [0,1)^2$, where every pure strategy δ_{x_i} of each player i is replaced by $\tilde{\delta}_{x_i} = (1-\beta_i)\delta_{x_i} + \beta_i\mathring{\phi}^{\bar{\kappa}}_i$; any mixed strategy $\mu_i \in M_i$ then becomes $\tilde{\mu}_i = (1-\beta_i)\mu_i + \beta_i\mathring{\phi}^{\bar{\kappa}}_i$ in $\Gamma_{\beta\mathring{\phi}^{\bar{\kappa}}}$.[8] By pre-stability of M, for every open Y containing M there must then exist a $\bar{\beta} > 0$ such that, for every β such that $\max_{i \in I} \beta_i \leq \bar{\beta}$, Y contains an equilibrium of $\Gamma_{\beta\mathring{\phi}^{\bar{\kappa}}}$. But I claim to the contrary that, for

[8]Note that this need not be the case for $\mu_i \notin M_i$, which may already satisfy the restriction $\beta_i\mathring{\phi}^{\bar{\kappa}}_i$.

any such Y and any $\tilde{\mu} \in Y$, there exists a deviation away from $\tilde{\mu}$ for each player i for any positive β_i. To see this, note that playing ϕ_i yields player i the same time-average payoff against μ_{-i} as μ_i for any $\mu \in M$, $\overline{U}_i(\phi_i, \mu_{-i}) = \overline{U}_i(\mu)$. Hence,

$$\begin{aligned}
\overline{U}_i(\tilde{\phi}_i, \tilde{\mu}_{-i}) &= (1-\beta_i)\,\overline{U}_i(\tilde{\phi}_i, \mu_{-i}) + \beta_i \overline{U}_i(\tilde{\phi}_i, \mathring{\phi}^{\bar{\kappa}}_{-i}) \\
&= (1-\beta_i)\left[(1-\beta_i)\overline{U}_i(\phi_i, \mu_{-i}) + \beta_i \overline{U}_i(\mathring{\phi}^{\bar{\kappa}}_i, \mu_{-i})\right] \\
&\qquad + \beta_i\left[(1-\beta_i)\overline{U}_i(\phi_i, \mathring{\phi}^{\bar{\kappa}}_{-i}) + \beta_i \overline{U}_i(\mathring{\phi}^{\bar{\kappa}}_i, \mathring{\phi}^{\bar{\kappa}}_{-i})\right] \\
&> (1-\beta_i)\left[(1-\beta_i)\overline{U}_i(\mu) + \beta_i \overline{U}_i(\mathring{\phi}^{\bar{\kappa}}_i, \mu_{-i})\right] \\
&\qquad + \beta_i\left[(1-\beta_i)\overline{U}_i(\mu_i, \mathring{\phi}^{\bar{\kappa}}_{-i}) + \beta_i \overline{U}_i(\mathring{\phi}^{\bar{\kappa}}_i, \mathring{\phi}^{\bar{\kappa}}_{-i})\right] \\
&= (1-\beta_i)\,\overline{U}_i(\tilde{\mu}_i, \mu_{-i}) + \beta_i \overline{U}_i(\tilde{\mu}_i, \mathring{\phi}^{\bar{\kappa}}_{-i}) \\
&= \overline{U}_i(\tilde{\mu})
\end{aligned}$$

for all positive β_i. ■

Intuitively, for an inefficient stage-Nash equilibrium, stage Nash leaves nothing to lose from experimenting with Pareto-improving play and reverting to stage Nash if the experimentation is not reciprocated.[9] Why is strong Pareto dominance required? Because the strategy profile where one player already plays according to ϕ plays like stage Nash, and the other player must be given a strict incentive to switch.

I can in fact generalize Proposition 1 to apply to sets of strategies with an inefficient stage-Nash continuation, on or off the equilibrium path. This captures the instability of sets of strategies that are "unforgiving" in the sense of having (inefficient) stage-Nash continuations, such as "grim–trigger" in the Prisoners' Dilemma. Given histories $g, h \in \overline{H}$, let gh be their *concatenation*, in the sense that history g is followed by history h. Given a stage-Nash equilibrium $\pi^* \in \Delta(A)$ of γ and a strategy profile $\hat{\mu}^* \in \Phi$ of Γ, if there exists a history $\hat{h} \in \overline{H}$ such that $\hat{\mu}^*|_{\hat{h}}$ plays π^* following any (continued) history, then $\hat{\mu}^*$ will be said to be a *trigger-strategy profile* of Γ with *trigger history* $\hat{h}$.

Theorem 1 *Suppose that $\pi^* \in \Delta(A)$ is a weakly inefficient pure equilibrium of γ, that $\hat{\mu}^* \in \Phi$ is a trigger-strategy profile of Γ with trigger history $\hat{h} \in \overline{H}$, and that $M \subseteq \Phi$ is a set of strategy profiles that play like $\hat{\mu}^*$. Then M is not stable under time-average payoffs.*

[9] These experimenting strategy profiles are inspired by Fudenberg and Maskin's (1990) secret-handshake strategies, but they are not quite the same; in particular, they never switch irrevocably to the Pareto-improving action profile. This is important under stability, where the secret-handshake signal may be sent "by mistake" with positive probability in the perturbed game $\Gamma_{\beta\mathring{\phi}^{\bar{\kappa}}}$, in which case an irrevocable switch away from stage Nash may be worse for a player than reversion to stage Nash.

Proof. As in the proof of Proposition 1, let $\bar{p} \in \mathscr{P}$ be a play that yields the players a time-average payoff by which π^* is strongly dominated, and $\bar{x}$ be the pure-strategy profile that induces $\bar{p}$ following any history. Consider a strategy profile $\psi \in \Phi$ such that, for all t:

- $\psi|_{\hat{h}h^t} = \delta_{\bar{x}}$ if $h^t = (a^0(\bar{x}), a^1(\bar{x}), \ldots, a^t(\bar{x}))$; and
- $\psi|_{h^t} = \hat{\mu}^*|_{h^t}$ otherwise.

Let $\mathring{\nu} \in \Phi$ be such that $\mathring{\nu}|_h = \mathring{a}$ for all $h \in \overline{H}$, where $\mathring{a}$ is a full-support mixture over action profiles (i.e. $\mathring{a}$ puts strictly positive probability on each $a \in A$). Now consider a strategy profile $\mathring{\psi}^\kappa \in \Phi$ such that $\mathring{\psi}^\kappa = (1-\kappa)\psi + \kappa\mathring{\nu}$, where $\kappa \in (0,1)$. Then I claim that there exists $\bar{\kappa}$ sufficiently small that $\mathring{\psi}^{\bar{\kappa}}$ is a completely mixed-strategy profile with $\overline{U}_i(\psi_i, \mathring{\psi}^{\bar{\kappa}}_{-i}) > \overline{U}_i(\mu_i, \mathring{\psi}^{\bar{\kappa}}_{-i})$ for both i and all $\mu \in M$. To see this, note first that if $\mathring{\psi}^{\bar{\kappa}}$ were not completely mixed, there would exist an open set $E \subset X$ such that $\mathring{\psi}^{\bar{\kappa}}(E) = 0$, and hence some history $h \in \overline{H}$ with zero probability under $\mathring{\psi}^{\bar{\kappa}}$ (i.e. $\xi_{\mathring{\psi}^{\bar{\kappa}}}(C(h)) = 0$), contradicting $\mathring{a}$'s full support. Second, clearly $\overline{U}_i(\psi|_{\hat{h}}) > \overline{U}_i(\mu_i|_{\hat{h}}, \psi_{-i}|_{\hat{h}})$ for all $\mu \in M$. Third, because $\mathring{a}$ is a full-support action-profile mixture (so that $(\psi_i|_h, \mathring{\nu}_{-i}|_h)$ will play like $(\mu_i|_h, \mathring{\nu}_{-i}|_h)$), $\overline{U}_i(\psi_i|_h, \mathring{\nu}_{-i}|_h) = \overline{U}_i(\mu_i|_h, \mathring{\nu}_{-i}|_h)$ for all $\mu \in M$ and all $h \in \overline{H}$; hence, $\overline{U}_i(\mathring{\psi}^{\bar{\kappa}}_i|_h, \mathring{\nu}_{-i}|_h) = \overline{U}_i(\mathring{\mu}^{\bar{\kappa}}_i|_h, \mathring{\nu}_{-i}|_h)$, where $\mathring{\mu}^{\bar{\kappa}} = (1-\bar{\kappa})\mu + \bar{\kappa}\mathring{\nu}$. It follows that

$$\begin{aligned}
\overline{U}_i(\mathring{\psi}^{\bar{\kappa}}|_{\hat{h}}) &= (1-\bar{\kappa})^2\overline{U}_i(\psi|_{\hat{h}}) + \bar{\kappa}(1-\bar{\kappa})\overline{U}_i(\mathring{\nu}_i|_{\hat{h}}, \psi_{-i}|_{\hat{h}}) + (1-\bar{\kappa})\bar{\kappa}\overline{U}_i(\psi_i|_{\hat{h}}, \mathring{\nu}_{-i}|_{\hat{h}}) + \bar{\kappa}^2\overline{U}_i(\mathring{\nu}_{\hat{h}}) \\
&> (1-\bar{\kappa})^2\overline{U}_i(\mu_i|_{\hat{h}}, \psi_{-i}|_{\hat{h}}) + \bar{\kappa}(1-\bar{\kappa})\overline{U}_i(\mathring{\nu}_i|_{\hat{h}}, \psi_{-i}|_{\hat{h}}) + (1-\bar{\kappa})\bar{\kappa}\overline{U}_i(\mu_i|_{\hat{h}}, \mathring{\nu}_{-i}|_{\hat{h}}) + \bar{\kappa}^2\overline{U}_i(\mathring{\nu}_{\hat{h}}) \\
&= \overline{U}_i(\mathring{\mu}^{\bar{\kappa}}_i|_{\hat{h}}, \mathring{\psi}^{\bar{\kappa}}_{-i}|_{\hat{h}}),
\end{aligned}$$

and since $\hat{h}$ has strictly positive probability under each of $\mathring{\psi}^{\bar{\kappa}}$ and $(\mathring{\mu}^{\bar{\kappa}}_i, \mathring{\psi}^{\bar{\kappa}}_{-i})$ by Lemma 1, $\overline{U}_i(\mathring{\psi}^{\bar{\kappa}}) > \overline{U}_i(\mathring{\mu}^{\bar{\kappa}}_i, \mathring{\psi}^{\bar{\kappa}}_{-i})$. Fourth, $\overline{U}_i(\mathring{\psi}^{\bar{\kappa}}_i, \cdot) \to \overline{U}_i(\psi_i, \cdot)$ and $\overline{U}_i(\mathring{\mu}_i, \cdot) \to \overline{U}_i(\mu_i, \cdot)$ as $\bar{\kappa} \to 0$.

Next consider the perturbed game $\Gamma_{\beta\mathring{\psi}^{\bar{\kappa}}}$, $\beta \equiv (\beta_1, \beta_2)' \in [0,1)^2$, where every pure strategy δ_{x_i} of each player i is replaced by $\tilde{\delta}_{x_i} = (1-\beta_i)\delta_{x_i} + \beta_i\mathring{\psi}^{\bar{\kappa}}_i$. By pre-stability of M, for every open Y containing M there must then exist a $\bar{\beta} > 0$ such that, for every β such that $\max_{i\in I}\beta_i \leq \bar{\beta}$, Y contains an equilibrium of $\Gamma_{\beta\mathring{\psi}^{\bar{\kappa}}}$. But the trigger history $\hat{h}$ has positive probability in $\Gamma_{\beta\mathring{\psi}^{\bar{\kappa}}}$ for any $\beta \in (0,1)^2$ by Lemma 1, and the proof of Proposition 1 establishes that $\psi|_{\hat{h}}$ (and hence ψ) is a profitable deviation away from any $\tilde{\mu}|_{\hat{h}}$ (and hence $\tilde{\mu}$) such that $\tilde{\mu} \in Y$. ■

Intuitively, ψ is just like ϕ in the proof of Proposition 1, except that it waits until the trigger history $\hat{h}$ before commencing its experimentation.[10] At this point, there is again nothing to lose from experimenting with Pareto-improving play and reverting to stage Nash

[10]Of course, Proposition 1 is just the special case of this result with $\hat{h} = \emptyset$.

if the experimentation is not reciprocated. And since every history has positive probability under a completely mixed-strategy profile, it does not matter if such a continuation is off the equilibrium path; the perturbations to which stable sets must be robust lead such continuations to be possible.

4 (Y, η)-instability with Patient Players

An obvious question to arise at this point is whether these results carry over to the discounting case. In fact, for a given discount factor $\rho < 1$, the destabilizing experimentation of the previous section breaks down, because there is now a utility cost to departing from Nash play temporarily before reverting. However, there remains a sense in which stage Nash play is unstable as the players become arbitrarily patient, as shown by the following result.

Proposition 2 *Suppose that $\pi^* \in \Delta(A)$ is a weakly inefficient pure equilibrium of γ, that $\hat{\mu}^* \in \Phi$ is a trigger-strategy profile of Γ with trigger history $\hat{h} \in \overline{H}$, and that $M \subseteq \Phi$ is a set of strategy profiles that play like $\hat{\mu}^*$. Then, there exists a neighborhood $Y \subset \Phi$ of M such that, for any $\bar{\eta} \in \mathring{\mathscr{M}}$, there exists $\rho \in [0, 1)$ sufficiently large that M is not $(Y, \bar{\eta})$-pre-stable under average discounted payoffs.*

Proof. Suppose otherwise; then for any neighborhood $Y \subset \Phi$ of M, there exists an $\bar{\eta} \in \mathring{\mathscr{M}}$ such that, for all $\rho \in [0, 1)$, M is $(Y, \bar{\eta})$-pre-stable under average discounted payoffs. Now consider the perturbed game $\Gamma_{\beta \mathring{\phi}^{\bar{\kappa}}}$ from the proof of Proposition 1. By $(Y, \bar{\eta})$-pre-stability of M, there exists a $\bar{\beta} > 0$ such that, for every β with $\max_{i \in I} \beta_i \leq \bar{\beta}$, Y contains an equilibrium of $\Gamma_{\beta \mathring{\phi}^{\bar{\kappa}}}$. But for any $\mu \in M$, $U_i(\phi_i, \mu_{-i}) \to U_i(\mu)$ as $\rho \to 1$, and $U_i(\phi) > U_i(\mu_i, \phi_{-i})$ for any $\rho \in [0, 1)$. Hence, there exists a $\rho \in [0, 1)$ sufficiently large that $U_i(\tilde{\phi}_i, \tilde{\mu}_{-i}) > U_i(\tilde{\mu})$ for all $\beta_i \in (0, \bar{\beta}]$. Hence, for ρ sufficiently large, there exists a deviation away from any $\tilde{\mu} \in Y$ for each player i and any $\beta_i \in (0, \bar{\beta}]$. Extending the argument as in the proof of Theorem 1, a contradiction is reached. ■

In words, there exists a neighborhood of a set of trigger-strategy profiles such that we can always find a small perturbation of the game and a sufficiently high level of patience amongst the players that there is no equilibrium in that neighborhood. This is a weaker notion of instability than if, fixing players to be sufficiently patient, there existed a neighborhood of a set of trigger-strategy profiles such that we could always find a small perturbation of the game with no equilibrium in that neighborhood. This stronger instability does not hold: fixing ρ, there exists a neighborhood Y of a set of strategies that plays like stage

Nash such that, under any sufficiently small perturbation of the game reciprocal experimentation is too unlikely to offer a profitable deviation to a strategy profile outside of Y. Experimentation can be delayed, in order to reduce the discounted value of its utility cost in the event of failure, but this brings it closer to the original strategy in the metric d, and hence within Y under a sufficiently small perturbation. As a result, the weaker form of instability captured in Proposition 2 requires the players' degree of patience to depend on the size of perturbation to the game.

5 A Stable Set with Arbitrary Patience

But then what would constitute a stable set? Indeed, since we have no existence result under time-average payoffs, will any strategy profiles form a stable set? To answer this question, I will use the following concepts, which modify those of Swinkels (1992) to the current setting.

Definition 2 *Let $M \subseteq \Phi$, and let Y be a neighborhood of M. A* directional retract for M and Y *is a map $R : \Phi \to Y$ such that:*

1. *R is continuous;*
2. *$R(\tilde{\mu}) = \tilde{\mu}$, $\forall \tilde{\mu} \in Y$;*
3. *$R^{-1}(M) = M$; and*
4. *$\forall \sigma \in \Phi \backslash Y$, $\exists \lambda \in (0,1)$ and $\mu \in M$ such that $R(\sigma) = (1-\lambda)\mu + \lambda\sigma$.*

Definition 3 *Given a status quo strategy profile $\mu \in \Phi$, an entrant $\sigma \in \Phi$, and a post-entry population $\tilde{\mu} = (1-\lambda)\mu + \lambda\sigma$, $\lambda \in [0,1]$, I will say that σ is an* α-perfect entrant taking the population from μ to $\tilde{\mu}$ *if:*

1. *σ is completely mixed;*
2. *$\sigma \in \overline{BR}(\tilde{\mu}, \alpha\bar{\eta}^{\alpha})$ for some $\bar{\eta}^{\alpha} \in \mathring{\mathscr{M}}$; and*
3. *for all $i \in I$, and any Borel set $B_i \subset X_i$, if $\delta_{x_i} \notin \overline{BR}_i(\tilde{\mu})$ for all $x_i \in B_i$, then $\sigma_i(B_i) \leq \alpha$.*

I say that σ is a perfect entrant taking the population from μ to $\tilde{\mu}$ *if there is $\{(\mu^{\alpha}, \sigma^{\alpha}, \tilde{\mu}^{\alpha})\}_{\alpha \downarrow 0}$, $(\mu^{\alpha}, \sigma^{\alpha}, \tilde{\mu}^{\alpha}) \to (\mu, \sigma, \tilde{\mu})$ such that for each α, σ^{α} is an α-perfect entrant taking the population from μ^{α} to $\tilde{\mu}^{\alpha}$. A closed set $M \subseteq \Phi$ is* uniformly robust to perfect entrants (URPE) *if there is a neighborhood* $Y \subseteq \Phi$ *of* M *such that for all* $\mu \in M$, *if* σ *is a perfect entrant taking the population from* $\mu \in M$ *to* $\tilde{\mu} \in Y$, *then* $\tilde{\mu} \in M$. *If there is a directional retract for* M *and* Y, *then we say that* M admits a directional retract.

Note that condition 2 in Definition 3 modifies the corresponding condition of Swinkels (1992) by explicitly requiring that σ be a (restricted) best reply in the perturbed game $\Gamma_{\alpha\bar{\eta}^\alpha}$, where $\alpha\bar{\eta}^\alpha$ is some strictly positive measure on X of norm no greater than α. This addition is quite consistent with the standard notion of perfectness (and arguably implicit in Swinkels' formulation), and moreover is crucial for Theorem 2 below.

Clearly URPE sets exist, although they need not be subsets of Nash equilibria—the whole strategy space, for instance. Moreover, if a URPE set M is convex, then a directional retract may be constructed as follows. Choose any $\varepsilon > 0$ such that M is URPE for $Y \equiv \{\mu' \in \Phi \mid d(\mu', M) \leq \varepsilon\}$. For $\mu \in \Phi$, define $V(\mu) \equiv \arg\min_{\mu' \in M} d(\mu, \mu')$. By contrast with Swinkels (1992, p. 337), $V(\mu)$ is not uniquely defined, for the following reason.

Lemma 2 *The metric* d *is convex, but not strictly convex.*

Proof. Since

$$\begin{aligned}
d(\mu, \lambda\mu' + (1-\lambda)\mu'') &= \sum_{h \in \overline{H}} \frac{1 - \mathbf{P}_h(\mu, \lambda\mu' + (1-\lambda)\mu'')}{3^m} \\
&= \sum_{h \in \overline{H}} \frac{1 - \lambda \sum_{x \in X} \mu(x|h)\mu'(x|h) - (1-\lambda) \sum_{x \in X} \mu(x|h)\mu''(x|h)}{3^m} \\
&= \sum_{h \in \overline{H}} \frac{\lambda(1 - \mathbf{P}_h(\mu, \mu')) + (1-\lambda)(1 - \mathbf{P}_h(\mu, \mu''))}{3^m} \\
&= \lambda d(\mu, \mu') + (1-\lambda) d(\mu, \mu''),
\end{aligned}$$

the result follows. ■

Hence, V is not a continuous function, but rather a union of continuous functions and hence a lower hemicontinuous correspondence. Because Φ is compact and V is lower hemicontinuous with nonempty closed convex values, there exists a continuous selection v from V by the Michael Selection Theorem. Let $L(\mu) \equiv \{\tilde{\mu} \mid \tilde{\mu} \in Y, \exists \alpha \in [0,1] \text{ such that } \tilde{\mu} = (1-\alpha)v(\mu) + \alpha\mu\}$. Because $L(\mu)$ is continuous and convex-valued as a correspondence, $W(\mu) \equiv \arg\min_{\tilde{\mu} \in L(\mu)} d(\mu, \tilde{\mu})$ is again a union of continuous functions, and has nonempty closed convex values. Hence, we can again apply the Michael Selection Theorem to obtain the required continuous selection from W. Conditions 2–4 of Definition 2 are clear by construction.

The following result extends Swinkels' (1992) Theorem 2 to an infinite strategy space; much of the proof simply reproduces that of Swinkels for convenience. Intuitively, for any given perturbed game featuring in the definition of a stable set, the proof constructs a perfect entrant which is then shown to be an equilibrium of the perturbed game. Whilst the notation of time-average payoffs is used in the proof of this result, it would also hold under average discounted payoffs, and indeed in a general n-player infinite normal-form game.

Lemma 3 *Every URPE set M that admits a directional retract contains a stable set.*

Proof. For both i, there exists a homeomorphism f_i from X_i to the unit interval $[0,1]$, on which we may define the Lebesgue measure $\mathscr{L}$. Fix a neighborhood Y of M and directional retract R. Let $\{\eta^k\}_{k\in\mathbb{N}}$ be a sequence of measures in $\mathscr{M}$ converging to 0 as $k\to\infty$. For any $x_i\in X_i$, define the mixed strategy $\breve{x}_i^k \equiv (1-\eta_i^k(X_i))\delta_{x_i}+\eta_i^k$ that is closest to δ_{x_i} satisfying the η^k restriction; for any $\nu_i\in\Phi_i$, define $\breve{\nu}_i^k\equiv(1-\eta_i^k(X_i))\nu_i+\eta_i^k$. For $\alpha\in(0,1)$ and $k\in\mathbb{N}$, define

$$P_\alpha^k(\mu)=\left\{\sigma\;\middle|\;\begin{array}{l}(1)\ \sigma_i(B_i)\geq\frac{\alpha}{2}\mathscr{L}(f_i(B_i))\text{ for all }i\in I,\text{ any Borel set }B_i\subset X_i\\(2)\ \left\{\forall x_i\in B_i:\breve{x}_i\notin\overline{BR}_i(\mu,\eta^k)\right\}\Rightarrow\sigma_i(B_i)\leq\alpha,\ \forall i\in I,\forall B_i\subset X_i\end{array}\right\}.$$

Let $C_\alpha^k(\cdot)\equiv P_\alpha^k(R(\cdot))$. Because P_α^k is an upper hemicontinuous, nonempty, compact- and convex-valued correspondence, R is a continuous function, and Φ is a locally convex Hausdorff space, C_α^k has a fixed point σ_α^k for each α and k by the Kakutani–Glicksberg–Fan Theorem. There then exists some $\bar{\eta}_\alpha^k\in\mathring{\mathscr{M}}$ such that $\sigma_\alpha^k\in\overline{BR}(\sigma_\alpha^k,\alpha\bar{\eta}_\alpha^k)$. For any $\nu\in\Phi$, let $F(\nu)\in M$ be such that $R(\nu)$ is a convex combination of ν and $F(\nu)$. For $\nu\in\Phi\backslash Y$, such an $F(\nu)$ exists by Definition 2.4. For $\nu\in Y$, $\nu=R(\nu)$, and so any element of M will do (since ν is expressible as a convex combination of any element of Φ and itself). Then, σ_α^k is an α-perfect entrant in Γ_{η^k} taking the population from $F(\sigma_\alpha^k)\in M$ to $R(\sigma_\alpha^k)$.

For each k, choose a subsequence of $\{\sigma_\alpha^k\}_{\alpha\downarrow 0}$ such that σ_α^k converges to some σ^k. Choose a convergent subsequence of $\{\sigma^k\}_{k\in\mathbb{N}}$ with limit σ. I will now show that σ is a perfect entrant taking the population from some $\mu\in M$ to $R(\sigma)$. Choosing $\alpha>0$, there exists k such that $d(\sigma,\sigma^k)<\alpha/3$ and $d(\delta_x,\breve{x})<\alpha/3$, where $\breve{x}$ is $\breve{x}^k$ for this value of k. Using the continuity of R, k can be chosen such that in addition $d(R(\sigma),R(\sigma^k))<\alpha/3$. Let ζ^α be an $\alpha/2$-perfect entrant taking the population from $F(\zeta^\alpha)$ to $R(\zeta^\alpha)$ in Γ_{η^k} such that $d(\sigma^k,\zeta^\alpha)<\alpha/3$ and $d(R(\sigma^k),R(\zeta^\alpha))<\alpha/3$. Since there is a convergent subsequence of $\{\sigma_\alpha^k\}_{\alpha\downarrow 0}$ with limit σ^k, since R is continuous, and since an α-perfect entrant is also an α'-perfect entrant for any

$\alpha' > \alpha$, such an entrant exists. Since $d(R(\zeta^\alpha), R(\breve{\zeta}^\alpha)) < \alpha/3$,

$$d(R(\sigma), R(\breve{\zeta}^\alpha)) \leq d(R(\sigma), R(\sigma^k)) + d(R(\sigma^k), R(\zeta^\alpha)) + d(R(\zeta^\alpha), R(\breve{\zeta}^\alpha)) < \alpha. \tag{1}$$

Now, assume that $\delta_{x_i} \notin \overline{BR}_i(R(\breve{\zeta}^\alpha))$ for all x_i in some Borel set $B_i \subset X_i$. Then, $\breve{x}_i \notin \overline{BR}_i(R(\zeta^\alpha), \eta^k)$, and so $\zeta_i^\alpha(B_i) \leq \alpha/2$. But,

$$\breve{\zeta}_i^\alpha(B_i) = (1 - \eta_i^k(X_i))\zeta_i^\alpha(B_i) + \eta_i^k \leq (1 - \eta_i^k(X_i))\frac{\alpha}{2} + \eta_i^k,$$

and so as $\eta_i^k < d(\delta_x, \breve{x}) < \alpha/3 < \alpha/2$, $\breve{\zeta}_i^\alpha(B_i) < \alpha$. Thus, $\breve{\zeta}^\alpha$ is an α-perfect entrant in Γ taking the population from $F(\breve{\zeta}^\alpha)$ to $R(\breve{\zeta}^\alpha)$.

Take a sequence $\alpha \to 0$. By (1), $R(\breve{\zeta}^\alpha)$ converges to $R(\sigma)$. By construction, $d(\sigma, \zeta^\alpha) \leq d(\sigma, \sigma^k) + d(\sigma^k, \zeta^\alpha) < 2\alpha/3$, and so ζ^α converges to σ. Take a subsequence such that $F(\zeta^\alpha)$ also converges. Now, by definition of F, each $F(\zeta^\alpha) \in M$ and so as M is closed, $\lim_{\alpha\downarrow 0} F(\zeta^\alpha) \in M$. As $\alpha \to 0$, $\eta^k \to 0$, and so $d(\nu, \breve{\nu}) \to 0$ for any $\nu \in \Phi$. Thus, $\lim_{\alpha\downarrow 0} F(\zeta^\alpha) = \lim_{\alpha\downarrow 0} F(\breve{\zeta}^\alpha)$, $\lim_{\alpha\downarrow 0} R(\zeta^\alpha) = \lim_{\alpha\downarrow 0} R(\breve{\zeta}^\alpha) = R(\sigma)$, and $\lim_{\alpha\downarrow 0} \zeta^\alpha = \lim_{\alpha\downarrow 0} \breve{\zeta}^\alpha = \sigma$. So, σ is a perfect entrant taking the population from $\lim_{\alpha\downarrow 0} F(\zeta^\alpha) \in M$ to $R(\sigma) \in Y$.

As M is URPE by hypothesis, $R(\sigma)$ must be an element of M. Hence, σ must be an element of M, and so $R(\sigma_\alpha^k) = \sigma_\alpha^k$ for σ_α^k sufficiently close to σ. Thus for k large, σ^k is a perfect (and so Nash) equilibrium of Γ_{η^k}. The existence of a minimal closed subset of M having the desired property is then given by the following argument. Let Ψ be the (non-empty) collection of pre-stable sets in Φ, partially ordered by (weak) set inclusion. By Hausdorff's Maximality Principle, Ψ contains a maximal nested sub-collection. Let $\Psi' \subseteq \Psi$ be such a sub-collection, and let $\widetilde{M}$ be the intersection of all sets M' for which $M' \in \Psi'$. Since each set M' is non-empty and compact, so is $\widetilde{M}$, by the Cantor Intersection Theorem. Since Ψ' is nested, for any open $Y \in \Phi$ containing $\widetilde{M}$, there must be an $M' \in \Psi'$ that is also contained in Y. It follows (by pre-stability of M') that there exists an $\bar{\eta} \in \mathring{\mathscr{M}}$ such that, for every $\eta \in \mathscr{M}$ with $\eta \leq \bar{\eta}$, $\overline{BR}(\cdot, \bar{\eta})$ has a fixed point μ in Y, i.e. $\widetilde{M}$ is (pre-)stable. ■

With this result in place, I can analyze the stability of specific strategy profiles of Γ. An obvious candidate for stability in the Prisoners' Dilemma is the pure-strategy profile known as *perfect tit-for-tat*, whereby each player cooperates if and only if both players played the same action in the previous period, and which is hence forgiving in the sense of having no stage-Nash continuations. Consider the class $\mathscr{G}^*$ of repeated games $\Gamma \in \mathscr{G}$ for which: the stage game g has a pure equilibrium $a^* \in A$ that is strongly dominated by the payoff $u(\bar{a})$ in another pure profile $\bar{a} \in A$; and the payoff $u(\bar{a})$ is achieved by no $\pi \in \Delta(A)\backslash\{\delta_{\bar{a}}\}$. Infinite repetition of the Prisoners' Dilemma in Figure 1 belongs to this class.

Definition 4 *Given $\Gamma \in \mathscr{G}^*$, the pure-strategy profile $\underline{x} \in X$ is called* generalized perfect tit-for-tat *if, for all t:*

- $a^t(\underline{x}) = \bar{a}$ *if* $a^{t-1} \in \{\bar{a}, a^*, \emptyset\}$*; and*
- $a^t(\underline{x}) = a^*$ *otherwise.*

Theorem 2 *Given $\Gamma \in \mathscr{G}^*$, if $\underline{M} \subseteq \Phi$ is the set of all convex combinations of pure-strategy profiles that play like generalized perfect tit-for-tat $\underline{x}$, then $\underline{M}$ contains a stable set under time-average payoffs.*

Proof. If σ is a perfect entrant taking the population from $\delta_{\underline{x}}$ to $\tilde{\delta}_{\underline{x}} = (1-\lambda)\delta_{\underline{x}} + \lambda\sigma$, $\lambda \in [0,1]$, then σ_i^α must be a best reply to $\tilde{\delta}^\alpha_{\underline{x}_{-i}}$ for both i in some sequence of perturbed games $\{\Gamma_{\alpha\bar{\eta}^\alpha}\}_{\alpha\downarrow 0}$ with each $\bar{\eta}^\alpha \in \mathring{\mathscr{M}}$. In particular, since all $h \in \overline{H}$ have positive probability under the completely mixed $(\sigma_i^\alpha, \tilde{\delta}^\alpha_{\underline{x}_{-i}})$ by Lemma 1, it follows that $\overline{U}_i(\sigma_i^\alpha|_h, \tilde{\delta}^\alpha_{\underline{x}_{-i}}|_h) \geq \overline{U}_i(\delta^\alpha_{\underline{x}_i}|_h, \tilde{\delta}^\alpha_{\underline{x}_{-i}}|_h)$ for all $h \in \overline{H}$, which holds if and only if

$$(1-\lambda)\overline{U}_i(\sigma_i^\alpha|_h, \delta^\alpha_{\underline{x}_{-i}}|_h) + \lambda\overline{U}_i(\sigma^\alpha|_h) \geq (1-\lambda)\overline{U}_i(\delta^\alpha_{\underline{x}}|_h) + \lambda\overline{U}_i(\delta^\alpha_{\underline{x}_i}|_h, \sigma^\alpha_{-i}|_h).$$

Since this must hold for all sufficiently small λ, it follows that $\overline{U}_i(\sigma_i^\alpha|_h, \delta^\alpha_{\underline{x}_{-i}}|_h) \geq \overline{U}_i(\delta^\alpha_{\underline{x}}|_h)$. But since, for all $h \in \overline{H}$, $\delta_{\underline{x}}|_h = \lim_{\alpha\to 0}\delta^\alpha_{\underline{x}}|_h$ has both players play $\bar{a}$ in all but finitely many periods, $(\sigma_i|_h, \delta_{\underline{x}_{-i}}|_h) = \lim_{\alpha\to 0}(\sigma_i^\alpha|_h, \delta^\alpha_{\underline{x}_{-i}}|_h)$ must do likewise, and hence $(\sigma_i, \delta_{\underline{x}_{-i}})$ must randomize over pure-strategy profiles that play like $\underline{x}$. This argument can be repeated with $(\sigma_i, \delta_{\underline{x}_{-i}})$ in place of $\delta_{\underline{x}}$, or indeed any convex combination of pure-strategy profiles that play like $\delta_{\underline{x}}$. Thus, $\tilde{\delta}_{\underline{x}} = (1-2\lambda)\delta_{\underline{x}} + \lambda(\sigma_i, \delta_{\underline{x}_{-i}}) + \lambda(\delta_{\underline{x}_i}, \sigma_{-i})$ is itself a convex combination of pure-strategy profiles that play like $\underline{x}$, i.e. it belongs to $\underline{M}$. $\underline{M}$ is thus URPE, hence it contains a stable set by Lemma 3. ■

Intuitively, generalized perfect tit-for-tat is efficient in any continuation, and punishes a deviation if and only if it does not play like it. Hence, strategy profiles that do play like perfect tit-for-tat are robust to noisy optimizing mutants—i.e. they constitute a URPE set, and thus contain a stable set. Swinkels' (1992) stronger concept of "uniform robustness against equilibrium entrants" (UREE) is unsuitable for this purpose, because it will admit any equilibrium alternative best reply. A UREE set in the Prisoners' Dilemma must hence include grim–trigger for instance, whereas URPE sets require entrants to perform well in the presence of mistakes. Sets that are "closed under better replies"—which also contain stable sets by Ritzberger and Weibull's (1995) Proposition 4—will admit *any* alternative best reply and are hence inadequate here *a fortiori*. Strategy profiles that play like generalized perfect tit-for-tat punish deviations (both on and off the equilibrium path), and hence any entrant that performs well in the presence of mistakes must likewise punish deviations. If we could not in this way exclude non-punishing alternative best replies (such as "always cooperate"), in turn we could not exclude inefficient alternative best replies to those strategies (such as "always defect"). Such "drift" would prevent us from having an efficient set closed under better replies (with a stable subset); such sets may be much bigger than URPE and stable sets in this setting.

5.1 Properties of the stable set

Theorem 2 establishes the existence of a stable set in the class $\mathscr{G}^*$ of repeated games. But what of the other properties of stable sets established by Al-Najjar (1995), admissibility, iterated dominance and forward induction?

An equilibrium $\bar{\mu}$ of Γ is *perfect* if it is the limit point of a sequence $\{\mu^k\}$ of equilibria of the perturbed games Γ_{η^k} for some sequence of perturbations $\{\eta^k\}$ in $\mathring{\mathscr{M}}$ converging to 0. A pure strategy $y_i \in X_i$ *dominates* the pure strategy $x_i \in X$ if $\overline{U}_i(\delta_{y_i}, \mu_{-i}) \geq \overline{U}_i(\delta_{x_i}, \mu_{-i})$ for all μ_{-i} and with strict inequality for at least one μ_{-i}. Following Al-Najjar (1995): a set $B_i \subset X_i$ is called *dominated* if every strategy in B_i is dominated (by possibly different strategies); a strategy $x_i \in X_i$ is *strongly dominated* if it has a dominated open neighborhood; and a mixed-strategy profile μ is *admissible* if, for every player i and every strongly dominated strategy x_i, $\mu_i(Y_i) = 0$ for every dominated open set Y_i containing x_i.

For any $x_i \in X_i$, the mixed strategy that is 'closest' to x_i satisfying the η restriction is $\breve{x}_i \equiv (1 - \eta_i(X_i))\delta_{x_i} + \eta_i$. The proofs of the following two lemmas are essentially those of Al-Najjar's Lemma A.1(iii) and (iv), but avoiding the need for continuous utility functions.

Lemma 4 *Fix the perturbation $\eta \in \mathscr{M}$ and the Borel set $B_i \subset X_i$. Suppose that μ is an equilibrium for Γ_η and that for all $x_i \in B_i$, we have $\overline{U}_i(\delta_{\breve{x}_i}, \mu_{-i}) < \overline{U}_i(\mu)$, then $\mu_i(B_i) = \eta_i(B_i)$.*

Proof. Suppose, by way of contradiction, that $b = \mu_i(B_i) - \eta_i(B_i) > 0$. Since μ is an equilibrium of Γ_η, there exists a $y_i \in X_i$ such that $\overline{U}_i(\mu) = \overline{U}_i(\delta_{\breve{y}_i}, \mu_{-i})$. It follows that $\overline{U}_i(\delta_{\breve{y}_i}, \mu_{-i}) > \overline{U}_i(\delta_{\breve{x}_i}, \mu_{-i})$ for all $x_i \in B_i$, and hence $\overline{U}_i(\delta_{y_i}, \mu_{-i}) > \overline{U}_i(\delta_{x_i}, \mu_{-i})$. Define the mixed strategy

$$\hat{\mu}_i = \mu_{i_{[B_i^c]}} + \eta_{i_{[B_i]}} + b\delta_{y_i}.$$

Since $\mu_{i_{[B_i^c]}} + \eta_{i_{[B_i]}} \geq \eta_i$ and $b > 0$, it follows that $\hat{\mu}_i \geq \eta_i$. Moreover,

$$\hat{\mu}_i - \mu_i = b\delta_{y_i} - (\mu_i - \eta_i)_{[B_i]},$$

and hence

$$\begin{aligned}\overline{U}_i(\hat{\mu}_i, \mu_{-i}) - \overline{U}_i(\mu) &= \overline{U}_i(b\delta_{y_i} - (\mu_i - \eta_i)_{[B_i]}, \mu_{-i}) \\ &= \int_{X_i} (\overline{U}_i(\delta_{y_i}, \mu_{-i}) - \overline{U}_i(\delta_{x_i}, \mu_{-i}))\, d(\mu_i - \eta_i)_{[B_i]} > 0,\end{aligned}$$

where the strict inequality follows from the fact that we have an integral of a strictly positive function evaluated using a positive, non-zero measure. Thus, μ_i is not a best reply against μ_{-i} in the perturbed game Γ_η, contradicting the assumption that μ is an equilibrium. ■

Lemma 5 *If $\eta_{-i} \in \mathring{\mathscr{M}}_{-i}$, μ is an equilibrium for Γ_η and x_i is a dominated strategy, then $\overline{U}_i(\delta_{\breve{x}_i}, \mu_{-i}) < \overline{U}_i(\mu)$.*

Proof. Suppose that y_i dominates x_i. Since μ_{-i} is completely mixed, it is clear that $\overline{U}_i(\delta_{y_i}, \mu_{-i}) > \overline{U}_i(\delta_{x_i}, \mu_{-i})$, and hence $\overline{U}_i(\delta_{\breve{y}_i}, \mu_{-i}) > \overline{U}_i(\delta_{\breve{x}_i}, \mu_{-i})$. Since μ_i is a best reply to μ_{-i} in the perturbed game Γ_η, we have $\overline{U}_i(\mu) \geq \overline{U}_i(\delta_{\breve{y}_i}, \mu_{-i}) > \overline{U}_i(\delta_{\breve{x}_i}, \mu_{-i})$ as required. ■

Proposition 3 below is essentially Al-Najjar's Proposition 4.1, adapted for discontinuous utility. The proof of his Proposition 4.2—that every equilibrium in a stable set is perfect—makes no use of continuous utility functions, and hence holds for time-average payoffs; I state it here as Proposition 4 without proof, for convenience.

Proposition 3 (Admissibility) *If $\bar{\mu}$ is a perfect equilibrium, then for every dominated open neighborhood Y_i, $\bar{\mu}_i(Y_i) = 0$. In particular, every perfect equilibrium is admissible.*

Proof. Since $\bar{\mu}$ is perfect, it is the limit point of a sequence $\{\mu^k\}$ of equilibria of the perturbed games Γ_{η^k} for some sequence of perturbations $\{\eta^k\}$ in $\mathring{\mathscr{M}}$ converging to 0. Suppose, contrariwise to the result, that $\bar{\mu}_i(Y_i) > 0$ for some dominated open neighborhood Y_i. As in the proof of Al-Najjar's Proposition 4.1, we know that $\liminf_{k\to\infty} \mu_i^k(Y_i) \geq \bar{\mu}_i(Y_i)$ and $\lim_{k\to\infty} \eta_i^k(Y_i) = 0$, and hence $\mu_i^k(Y_i) > \eta_i^k(Y_i)$ for all large enough k. Lemma 4 now shows that, for all large k, there must be some $\bar{x}_i^k \in Y_i$ such that $\overline{U}_i(\mu^k) = \overline{U}_i(\delta_{\bar{x}_i^k}, \mu_{-i}^k)$. But since such $\bar{x}_i^k$ is dominated by definition and since μ^k is completely mixed, Lemma 5 yields the opposite conclusion that $\overline{U}_i(\mu^k) > \overline{U}_i(\delta_{\bar{x}_i^k}, \mu_{-i}^k)$. ■

Proposition 4 (Al-Najjar 1995) *Every equilibrium in a stable set is perfect.*

Thus, we see that admissibility of stable sets readily carries over to the infinite case with time-average payoffs. By contrast, Al-Najjar's demonstration that stable sets of infinite games satisfy iterated dominance and forward induction relies crucially on the continuity of utility functions. However, the properties of URPE sets compensate for the discontinuity of time-average payoffs and allow me to prove the following.

Proposition 5 (Iterated Dominance) *Suppose that M is a URPE set of the game Γ and that $B_i \subset X_i$ is a set of dominated strategies. Then M contains a URPE set in the game obtained from Γ by eliminating B_i.*

Proof. For any α-perfect entrant σ^α taking the population from μ^α to $\tilde{\mu}^\alpha$, $\tilde{\mu}^\alpha$ is completely mixed. Hence, for any $x_i \in B_i$, $\delta_{x_i} \notin \overline{BR}_i(\tilde{\mu}^\alpha)$ and $\sigma_i^\alpha(B_i) \leq \alpha$. And since M cannot be dominated (by Lemma 3 and Propositions 3 and 4), there exists some set $M' \subseteq M$ such that $\lim_{\alpha\to 0} \tilde{\mu}^\alpha(B_i) = 0$ and $\lim_{\alpha\to 0} \tilde{\mu}^\alpha \in M'$ for all $\mu \in M'$. ■

Note that, by contrast with Al-Najjar's Proposition 4.3, only dominatedness and not strong dominatedness of B_i is required here, and moreover B_i is not required to be compact.

Given a set M of equilibria, recall that x_i is *an inferior response to M* if, for every equilibrium $\mu \in M$,

$$\overline{U}_i(\mu) > \overline{U}_i(\delta_{x_i}, \mu_{-i}). \tag{2}$$

In the absence of continuous utility functions, I must strengthen this concept slightly: Given a set M of equilibria, x_i is *a strongly inferior response to M* if there exists a neighborhood Y of M such that the inequality in (2) holds for every $\mu \in Y$. Since the following result makes use of this stronger notion of an inferior response, I call it "weak" forward induction.

Proposition 6 (Weak Forward Induction) *Suppose that M is a URPE set of equilibria of Γ, and that $B_i \subset X_i$ is a set of strongly inferior responses to M. Then M is also a URPE set in the game obtained from Γ by eliminating B_i.*

Proof. There exists a neighborhood $Y \subseteq \Phi$ of M such that, for any $x_i \in B_i$ and any perfect entrant σ taking the population from $\mu \in M$ to $\tilde{\mu} \in Y$, $\delta_{x_i} \notin \overline{BR}_i(\tilde{\mu}^\alpha)$ and hence $\sigma_i^\alpha(B_i) \leq \alpha$. Hence, $\tilde{\sigma}(B_i) = \lim_{\alpha\to 0} \tilde{\sigma}^\alpha(B_i) = 0$, and since any $\mu \in M$ is an equilibrium and $\overline{U}_i(\mu) > \overline{U}_i(\delta_{x_i}, \mu_{-i})$ for any $x_i \in B_i$, it follows that $\mu(B_i) = \tilde{\mu}(B_i) = 0$. ■

Therefore, Theorem 2's set $\underline{M}$ of all convex combinations of pure-strategy profiles that play like generalized perfect tit-for-tat satisfies iterated dominance and weak forward induction, whilst its stable subset is admissible.

6 Discussion

Which of these properties is responsible for the results of this paper? It is not admissibility or iterated dominance, for there are certainly strategies against which inefficient stage Nash would do better than experimentation with Pareto-improving play. Is it weak forward induction then? Ben-Porath and Dekel (1992) show that, in two-person games with common interests, only the optimal outcome survives iterated elimination of dominated strategies when one player is allowed to "burn money". Forward induction is also key to Osborne's (1990) near-optimality in finitely repeated coordination games, but his argument loses force when equilibrium paths contain outcomes which are not stage-Nash equilibria. Moreover, the experimentation strategies of this paper involve no explicit signalling. However, for Kohlberg and Mertens (1986, p. 1029), forward induction means that "a stable set contains a stable set of any game obtained by deletion of a strategy which is an inferior response in all equilibria of the set". And under time-average payoffs, inefficient stage Nash would be an inferior response (indeed, a strongly inferior response) to a set of strategy profiles that experimented with Pareto-improving play. Hence, it is (weak) forward induction that drives the instability of inefficient stage Nash, and the perturbations of strategic stability that extend this instability to inefficient stage-Nash continuations.

Finally, in no sense do the results of this paper constitute a refinement of the Folk Theorem. Indeed, I conjecture that any feasible, individually rational payoffs could be sustained by a strategically stable set, including for instance $(1,1)$ in the Prisoners' Dilemma of Figure 1. However, the strategy profiles involved would not include inefficient stage-Nash continuations. Moreover, they would have to be quite complex; an inefficient version of perfect tit-for-tat would not be stable, for instance, because there would be nothing to lose by experimenting with Pareto-improving play. To be stable in fact, inefficient strategy profiles would have to threaten to punish experimentation, and they would have to do so in any continuation. Hence, whilst equilibrium payoffs are not constrained by strategic stability, behavior off the equilibrium path is significantly constrained, and in particular we have a refinement that rules out the unforgiving play of inefficient stage Nash.

References

Al-Najjar, N. (1995): "Strategically Stable Equilibria in Games with Infinitely Many Pure Strategies," *Mathematical Social Sciences*, 29, 151–164.

Aliprantis, C. D., and K. C. Border (2007): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin, 3rd edn.

Aumann, R. J., and S. Sorin (1989): "Cooperation and Bounded Recall," *Games and Economic Behavior*, 1, 5–39.

Ausubel, L. M., and R. J. Deneckere (1993): "A Generalized Theorem of the Maximum," *Economic Theory*, 3, 99–107.

Axelrod, R. (1981): "The Emergence of Cooperation Among Egoists," *American Political Science Review*, 75, 306–318.

——— (1984): *The Evolution of Cooperation*. Basic Books, New York.

Axelrod, R., and W. Hamilton (1981): "The Evolution of Cooperation," *Science*, 211, 1390–1396.

Bagwell, K., and G. Ramey (1996): "Capacity, Entry, and Forward Induction," *RAND Journal of Economics*, 27, 660–680.

Ben-Porath, E., and E. Dekel (1992): "Signaling Future Actions and the Potential for Sacrifice," *Journal of Economic Theory*, 57, 36–51.

Binmore, K. G., and L. Samuelson (1992): "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory*, 57, 278–305.

Boyd, R. (1989): "Mistakes Allow Evolutionary Stability in the Repeated Prisoner's Dilemma Game," *Journal of Theoretical Biology*, 136, 47–56.

Boyd, R., and J. Lorberbaum (1987): "No Pure Strategy is Evolutionarily Stable in the Repeated Prisoners' Dilemma Game," *Nature*, 327, 58–59.

Carbonell-Nicolau, O. (2011): "On Strategic Stability in Discontinuous Games," *Economics Letters*, 113, 120–123.

Cho, I.-K., and D. M. Kreps (1987): "Signalling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179–221.

Cho, I.-K., and J. Sobel (1990): "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50, 381–413.

Demange, G. (1992): "Rational Escalation," *Annales d'Économie et de Statistique*, 25/26, 227–249.

Demichelis, S., and K. Ritzberger (2003): "From Evolutionary to Strategic Stability," *Journal of Economic Theory*, 113, 51–75.

Dixit, A. (1980): "The Role of Investment in Entry-Deterrence," *Economic Journal*, 90, 95–106.

Farrell, J., and R. Ware (1989): "Evolutionary Stability in the Repeated Prisoner's Dilemma," *Theoretical Population Biology*, 36, 161–166.

Friedman, J. W. (1971): "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies*, 38, 1–12.

Fudenberg, D., and E. Maskin (1990): "Evolution and Cooperation in Repeated Games," *American Economic Review Papers and Proceedings*, 80, 274–279.

Glazer, J., and A. Weiss (1990): "Pricing and Coordination: Strategically Stable Equilibria," *Games and Economic Behavior*, 2, 118–128.

Govindan, S. (1995): "Stability and the Chain Store Paradox," *Journal of Economic Theory*, 66, 536–547.

Hillas, J. (1990): "On the Definition of the Strategic Stability of Equilibria," *Econometrica*, 56, 1365–1391.

Kim, Y. (1994): "Evolutionary Stable Strategies in the Repeated Prisoner's Dilemma," *Mathematical Social Sciences*, 28, 167–197.

Kohlberg, E., and J.-F. Mertens (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003–1037.

Kreps, D. M., and R. Wilson (1982a): "Reputation and Information," *Journal of Economic Theory*, 27, 253–279.

——— (1982b): "Sequential Equilibria," *Econometrica*, 50, 863–894.

Leininger, W. (1984): "A Generalization of the 'Maximum Theorem'," *Economics Letters*, 15, 309–313.

Mertens, J.-F. (1989): "Stable Equilibria: A Reformulation Part I. Definition and Basic Properties," *Mathematics of Operations Research*, 14, 575–625.

——— (1991): "Stable Equilibria: A Reformulation Part II. Discussion of the Definition and Further Results," *Mathematics of Operations Research*, 16, 694–753.

Myerson, R. B. (1978): "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 73–80.

Osborne, M. J. (1990): "Signalling, Forward Induction, and Stability in Finitely Repeated Games," *Journal of Economic Theory*, 50, 22–36.

Ritzberger, K., and J. W. Weibull (1995): "Evolutionary Selection in Normal-Form Games," *Econometrica*, 63, 1371–1399.

Robson, A. J. (1990): "Efficiency in Evolutionary Games: Darwin, Nash and the Secret Handshake," *Journal of Theoretical Biology*, 144, 379–396.

Selten, R. (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 25–55.

——— (1983): "Evolutionary Stability in Extensive Two-Person Games," *Mathematical Social Sciences*, 5, 269–363.

Sugden, R. (1986): *The Economics of Rights, Cooperation and Welfare*. Basil Blackwell, Oxford.

Swinkels, J. M. (1992): "Evolution and Strategic Stability: From Maynard Smith to Kohlberg and Mertens," *Journal of Economic Theory*, 57, 333–342.

——— (1993): "Adjustment Dynamics and Rational Play in Games," *Games and Economic Behavior*, 5, 455–484.

Thomas, B. (1985): "On Evolutionarily Stable Sets," *Journal of Mathematical Biology*, 22, 105–115.

van Damme, E. (1984): "A Relation between Perfect Equilibria in Extensive Form Games and Proper Equilibria in Normal Form Games," *International Journal of Game Theory*, 13, 1–13.

——— (1987): *Stability and Perfection of Nash Equilibria*. Springer Verlag, Berlin.

——— (1989): "Stable Equilibria and Forward Induction," *Journal of Economic Theory*, 48, 476–496.