Repeated Signalling and Reputation

Charles Roddie*
Princeton University

August 2008
For the latest version visit: `www.princeton.edu/~croddie`

Abstract

The one-shot monotonic signalling game can be solved under the refinement D1, selecting the Riley equilibrium. This paper adapts the refinement to the repeated signalling game, selecting a dynamic version of the Riley equilibrium, defined iteratively, in which types minimally separate in each period.

This model provides an alternative framework for studying reputation, generating under appropriate limits a modified Stackelberg property: each type above the lowest takes the action that maximizes Stackelberg payoffs, subject to separating from the lowest type. In contrast to the usual approach to reputation there are no behavioural types. It can be solved under arbitrary discount factors of both players: if the signaller discounts, the result above holds with the signaller's Stackelberg payoffs replaced by simply defined "discounted Stackelberg" payoffs. If the respondent has preferences not only over the actions but also over the type of the signaller, a differential equation characterizes the limit, combining reputational and pure type-signalling motives.

1 Introduction

1.1 A signalling model of reputation

The economic idea of reputation is that a patient player by taking a certain action may cause others to expect him to do the same thing in the future, even if it will be against his immediate interests. By doing this he has effectively the ability to commit to any

*I am greatly indebted to Wolfgang Pesendorfer for insights, encouragement, criticism, and patience, to Stephen Morris for advice and detailed comments, and to Dilip Abreu for valuable suggestions. The paper has benefitted from discussions with Satoru Takahashi, and feedback from seminar participants at Princeton, LSE, NYU and Rochester.

action, receiving (close to) Stackelberg payoffs[1]. In the standard model of reputation ([11], [12], [17]) this conclusion requires the possibility of the patient player being a *behavioural* or *commitment* type. Such a type uses a fixed exogenous strategy, independent of expectations about the other player's strategy; in the simplest case this strategy is a particular action that is always taken. The normal type(s) of the long-run player can then develop a reputation for one of these actions over time by repeatedly mimicing one of these behavioural types, if his discount factor is high enough.

The signalling model of reputation proposed here has a different logic. Here there are only normal types of the signaller, giving a range of preferences. Instead of pooling with a behavioural type, each normal type separates from "worse" types. Types are correlated over time and this separation occurs at every stage. The signaller wants to be seen as a *higher* type and this give an incentive to take *higher* actions than are myopically optimal: by taking a *higher* action today, the signaller will be seen as a *higher* type today, and so will be expected expected to be a *higher* type tomorrow, and so expected to take a *higher* action tomorrow, leading to more favourable treatment by the other player. This is the reputational incentive to take higher actions than would be myopically optimal.

Suppose that types are unlikely to change from one period to the next (an assumption) and that each type's action does not change much from a given period to the next (a limit property as the length of the game tends to infinity). What generates a Stackelberg property is that by taking an action, the signaller signals that he is the type that preferred to take this action, and will be expected to be the same type in the following stage, and so to do (approximately) the same action in the next period. This holds only when the signaller chooses an action that is taken with positive probability by some type; the set of these such actions in the limit determines how the Stackelberg result is qualified.

Assume that the signaller is patient and that the receiver's preferences are over actions only and do not involve the signaller's type - as in standard reputation models. Taking the limit as the number of periods from the end of the game tends to infinity, as types become dense in some interval, and when type change becomes infinitely unlikely, we get the following reputation result: The lowest type takes his myopically optimal action. All other types take the actions that give them the greatest Stackelberg payoff, subject to separating (by the lowest type's Stackelberg payoff) from the lowest type[2]. So low types take the minimal action that separates them from the lowest type, while high types take their Stackelberg actions. The limit result is a combination of a separation property, such as is often seen in signalling models, with a Stackelberg property, often seen in reputation models.

The signalling model is more tractable than standard reputation models tend to be, generating a unique, simple and calculable solution under the refinement. In particular, the model remains solvable under general discount factors of both players. Standard reputational models require the reputation-builder to have discount factor 1 or tending

[1] The Stackelberg payoff is the payoff in the stage game to a player who can commit to any action, while the other player best-responds. A Stackelberg action maximizes this payoff.

[2] That is to say, they take actions that make the lowest type not want to mimic them.

to 1, and often myopic play by the other player, or at least a level of patience that becomes infinitely less than the reputation-builder's. In the repeated signalling model, the respondent's discount factor has no effect on the solution, and when the signaller has discount factor other than 1, the result given above only requires Stackelberg payoffs to be replaced with simply defined "discounted Stackelberg" payoffs.

The type of the signaller and actions of both players lie in intervals of real numbers, with monotonicity properties that will be spelled out presently. This is a loss of generality from the standard reputation model, which can work with very general stage games. But it is a natural specification for a large class of applied models. Examples include developing reputations for product quality, with the signaller's type being firm quality, monetary policy of a central bank, with type being toughness on inflation, and work incentives, with type being ability. As a model of work incentives it could be seen as a development of Holmstrom's model [14]. In that model the worker signals his ability, but without knowing his own ability - which makes strategies simpler. Fudenberg and Tirole [13] refer to this type of model as a "signal-jamming" model. A combination of signal, the signaller's action, and random noise is observed at each stage. In a particular specification of incentives, with work and ability being perfect substitutes, and with normal distributions of noise, there is a Stackelberg result. A repeated signalling model does the same type of thing with a more standard approach to signalling, and allowing for more general specifications of payoffs and information.

Mailath and Samuelson ([19], [20]) have also studied reputational issues in terms of dynamic signalling. In a two-type model they find that reputation effects can be supported with the high type separating from the low type. In common with the repeated signalling model, the effect is generated by separation from bad type(s) rather than pooling with a "good" type. The importance of types being changeable over time is also emphasized. There is no clear relation to Stackelberg actions and payoffs: their interest is in qualitatively supporting reputational concerns and also seeing how reputation can be built up and lost gradually.

1.2 From static to repeated signalling

Signalling games have been a fruitful area in pure and applied work, beginning with Spence's model of education and job-market signalling [24]. See Sobel [23] for a survey of applications. In the canonical monotonic signalling model the signaller has a type, which is private information, and takes an action in a space embedded in $\mathbb{R}$ or $\mathbb{R}^n$. Types are a subset (finite or continuum) of some real interval, and higher types have more of a preference for higher actions, a single-crossing condition. The respondent observes the action, forms a belief about the type, and replies accordingly, treating higher types more favourably - a preference given by external considerations. In Spence's example, types who find education easier are given a higher wage because they are expected to do better work: this is not modelled but is the reason for the respondent's preference.

In the repeated signalling model both players move simultaneously in the stage game[3] and as described above, which allows the study of reputation, and there are reputational reasons for higher types being treated more favourably. Beyond the usual assumptions of monotonic signalling games there is currently an additive separability assumption: the signaller's stage-game payoff is separable between the respondent's action and the signaller's type and action. This provides uniqueness of equilibrium in the stage game and allows simple characterization of the solution and allows discounted Stackelberg payoffs to be defined simply.[4]

There are many perfect Bayesian equilibria of the one-shot signalling game, some separating and some pooling, and the most used equilibrium of the signalling game is the Riley equilibrium, which is minimally separating. That is to say no two types take the same action with positive probability, and each type takes his most preferred action given the requirement of separating from lower types (so that no lower type would want to mimic him). In a finitely repeated signalling game, applying this property at every stage, starting with the last, gives what I call the iterated Riley equilibrium, in which types separate minimally at each stage. This means that at every stage the current type of the signaller is revealed. (Type is changeable but correlated over time, following a Markov process). This equilibrium has a particularly simple form, with the signaller's actions depending only on his current type and how many periods he is from the end of the game, and the respondent responding myopically (regardless of his actual discount factor) to current expectations of play. To calculate the signaller's strategy in a given period, we only need to find the minimally separating equilibrium between his current action and the respondent's response in the next period. This response depends on the signaller's strategy in the next period and it is an inductive calculation.

From a theoretical angle, attempts were made to cut down the number of equilibria with restrictions on beliefs off the equilibrium path, some beliefs being considered more reasonable than others. Particularly successful are the set of related refinements that go by the name "divinity", including D1, defined in Cho, Kreps [5].[5] In this paper I consider a refinement, labelled D_ω, which is an extension of divinity to the repeated signalling game. The spirit of the refinement D_ω is this: sub-optimal actions by the signaller are interpreted as over-confidence, over-confidence about the respondent's response to these actions. I define the justifying beliefs of an action to be those beliefs about the respondent's immediate response that would justify the action over the equilibrium action. And the criterion D_ω is that if one type has a smaller set of justifying beliefs for a particular action than another type, then the first type is ruled out, assuming that the second type was given positive probability before the action was observed. Suppose, informally, that larger belief-mistakes are infinitely less likely than smaller ones, uniformly across

[3] Although this has not been formally shown, there will be no change to any results if the respondent moves first.

[4] Work is in progress to relax this assumption.

[5] Other criteria less connected to the current work include the *intuitive criterion* of Cho, Kreps [5], which provides a unique solution when there are only two types, strategic stability [16], which is defined on finite action spaces, the weak condition of "*undefeated equilibium*" (Mailath, Okuno-Fujiwara, and Postlewaite [18]), and evolutionary stability ([21],[3]).

types. Then any type requiring a larger belief-mistake (overconfidence) to justify an action than another type must be assigned infinitely less probability than this other type. The criterion D_ω is a weakening of this condition.

Cho, Sobel [6] show that in monotonic signalling games the D1 criterion selects the Riley equilibrium uniquely (assuming pooling at the highest action is ruled out). A similar logic is used here to show that D_ω uniquely selects the iterated Riley equilibrium. There are two steps. First pooling is eliminated. At any point in the game if two types pool on the same action, it is shown that by taking a slightly higher action, the signaller is considered to be at least the higher type, so payoffs increase discontinuously on raising the action from this point. Second, separating is shown to be minimal. If a type does more takes a more costly action than necessary to separate from the preceding type then changing the action slightly is shown not to affect beliefs, so the original action cannot be optimal.

When type-change from period to period becomes very unlikely I find limit properties as the number of periods from the end of the game tends to infinity, calculating a limit map from types of the signaller to actions. When types become dense in an interval we get the modified Stackelberg result given above. This happens when the respondent has preferences over actions of both players, and not over the type of the signaller, so that signalling incentives derive entirely from reputational concerns. When the respondent has preferences over actions and the signaller's type, rewarding both high expected actions and high types, I find that the limit map from types to actions is characterized by a differential equation. This is the same differential equation as in the Riley equilibrium of the single-stage game when the signaller moves first, so is already a Stackelberg leader, with a different starting point. Thus the limit combines commitment and pure type-signalling motives, commitment motives arising from reputational considerations.

1.3 Contents

Chapter 2 defines the model and states its main assumptions. Chapter 3 defines perfect-Bayesian equilibrium and the D_ω refinement. I define the iterated Riley solution in chapter 4, and show how it results uniquely from the D_ω refinement in chapter 5. In chapter 6 limit properties are found as type change becomes very unlikely and as the number of periods tends to infinity. Limit properties of the reputation case are derived and discussed in chapter 7, and limits of the general case in chapter 8. Chapter 9 proposes further work.

2 Model

2.1 Actions, types and utilities

There are two players and k periods. In each period both players take actions simultaneously; actions are observable. The signaller takes actions from the set $A \subseteq \mathbb{R}$, $A = [a_{\min}, a_{\max}]$; the respondent takes actions from $R \subseteq \mathbb{R}$, $R = [r_{\min}, r_{\max}]$.

The respondent has no private information; he has a type in each period (a "period-type") which determines both players' payoff functions in the stage game in that period. The signaller in each period knows his current and previous period-types. Each period-type lies in a finite set $T \subseteq \mathbb{R}$, $T = \{\tau_0, ...\tau_h\}$. Let the global type, the vector of period-types, of the signaller be $t^k \in T^k$; the signaller's period-type in period i is then t_i. A sub-vector of types $t^n \in T^n$ describes period-types in periods 1 to n.

The signaller has the discounted utility function $U_1 = \sum_i \delta_1^i u_1 (t_i, a_i, r_i)$ from outcomes $O = (T \times A \times R)^k$ to $\mathbb{R}$, with u_1 a continuous function $T \times A \times R \to \mathbb{R}$ and with $0 < \delta_1 \leq 1$.

The respondent has utility function $U_2 = \sum_i \delta_2^i u_2 (t_i, a_i, r_i) : O \to \mathbb{R}$, with u_2 a continuous function $T \times A \times R \to \mathbb{R}$ and with $0 < d_2$.

2.1.1 Assumptions on u_2

Assumption 1 *$\int u_2(., r) d\mu_{ta}$ is strictly quasi-concave in r for any probability measure μ_{ta} on $T \times A$.*

The above integral is continuous by continuity of u_2 and so has a maximum in r for each probability measure μ_{ta}. The quasi-concavity assumption ensures that there is a unique maximum. Call this maximum $r^* (\mu_{ta})$, the myopic best response of the respondent to the belief μ_{ta}.

Definition 1 *For any measure μ_{ta} on $T \times A$, let $r^* (\mu_{ta}) = \arg\max \int u_2(., r) d\mu_{ta}$*

Assumption 2 $\mathrm{Im}(r^*) \subseteq (r_{\min}, r_{\max})$

Assumption 3 *Increasing response to types or actions:*

u_2 is differentiable in the third argument and $(\partial/\partial r) u_2 (t, a, r)$ is strictly increasing in (t, a).

Here $(t_1, a_1) < (t_2, a_2)$ iff $t_1 \leq t_2$ and $a_1 \leq a_2$ with at least one inequality strict.

Assumptions 2 and 3 imply the following fact, which is their only role in this paper: fixing a map between types and actions, if the distribution of types increases in the sense of first order stochastic dominance, then the myopic best response of the respondent will increase.

Fact 1 *If $\alpha : T \to A$ is a strictly increasing function, and $f : T \to T \times A$ by $f(\tau) = (\tau, \alpha(\tau))$, and if $\mu_t < \mu'_t$ in the sense of first order stochastic dominance, then $r^*(\mu_{ta}) < r^*(\mu'_{ta})$, where $\mu_{ta} = \mu_t \circ f^{-1}$ and $\mu'_{ta} = \mu'_t \circ f^{-1}$.*

Proof. $r^*(\mu_{ta})$ maximizes the quasi-concave function of r, $\int u_2(t,a,r)\, d(\mu_t \circ f^{-1})(t,a)$, so $(\partial/\partial r)\int u_2(t,a,r)\, d(\mu_t \circ f^{-1})(t,a) = \int (\partial/\partial r) u_2(t,a,r)\, d(\mu_t \circ f^{-1})(t,a) = 0$ at $r^*(\mu_{ta})$, so $\int (\partial/\partial r) u_2(t,a,r)\, d(\mu'_t \circ f^{-1})(t,a) > 0$ at $r^*(\mu_{ta})$, but $= 0$ at $r^*(\mu_{ta})$, so we must have $r^*(\mu_{ta}) < r^*(\mu'_{ta})$ by quasi-concavity of $\int u_2(t,a,r)\, d(\mu'_t \circ f^{-1})(t,a)$. ■

Say that any function $\widehat{r} : \Delta(T \times A) \to R$ for which the above property holds satisfies *increasing response to types or actions.*

2.1.2 Assumptions on u_1

Assumption 4 *Additive separability: $u_1(t,a,r) \equiv v_a(t,a) + v_r(r)$*

Assumption 5 *v_r is strictly increasing*

The first monotonicity assumption for the signaller, assumption 5 requires that higher actions by the respondent are preferred by the signaller. It is equivalent to $u_1(t,a,.)$ being a strictly increasing function for each t, a.

Assumption 6 *Single crossing: If $a_1 < a_2$ and $t_1 < t_2$ then $v_a(t_2,a_2) - v_a(t_2,a_1) > v_a(t_1,a_2) - v_a(t_1,a_1)$*

The second monotonicity assumption, what assumption 6 expresses is that higher types of the signaller are more disposed to taking higher actions.

It follows from this assumption that if $a_1 < a_2$ and $t_1 < t_2$ then $u_1(t_1,a_1,r_1) \leq u_1(t_1,a_2,r_2)$ implies $u_1(t_2,a_1,r_1) < u_1(t_2,a_2,r_2)$; and this is condition that will be used in this paper

Assumption 7 *$v_a(t,a)$ is strictly quasi-concave in a for all t.*

This is equivalent to $u_1(t,a,r)$ being strictly quasi-concave in a for all t and r. This assumption will be used to give unique solutions to optimization problems by the signaller. In particular, fixing r, there is a be a unique action for any type which maximizes $v_a(t,a)$: this is implied by quasi-concavity and continuity of v_a. We will call this action $a^*(t)$.

Definition 2 $a^*(t) = \arg\max_a v_a(t, a)$

Assumption 8 *Undesirable $a_{\min}$ and $a_{\max}$:*

For each $t \in T$, $t > \tau_0$, $v_a(t, a)$ is not maximized at $a_{\min}$.

For any $r_1, r_2 \in \mathrm{Im}(r^*)$, $t \in T$, $v_a(t, a^*(t)) + v_r(r_1) > v_a(t, a_{\max}) + v_r(r_2)$

The undesirable $a_{\max}$ assumption is that no change in the respondent's action (within the myopic best-response set $\mathrm{Im}(r^*)$) will compensate for taking the action $a_{\max}$ over $a^*(t)$.

It is important to eliminate the possibility of pooling at the highest action because then my game structure in which types are revealed each period breaks down. Also, without pooling there will be a simple map from types to actions in a given period, independent of the current type distribution, while the type-action correspondence for a pooled equilibrium depends on the type distribution.

The undesirable $a_{\min}$ assumption is needed to ensure that the map from types to actions is always strictly increasing, required to generate the limit reputation result. If the respondent has a concern for type as well as actions then we do not need this assumption as the map will be strictly increasing from the penultimate stage back. (It is possible that a reworking of the limit results will eliminate the need for either of these alternatives.)

Fact 2 *If the above assumptions 4-8 on u_1 are satisfied by $u_1 = v_a(t, a) + v_r(r)$ then they are satisfied by $u^E = v_a(t, a) + \delta_1 v_r(r)$ for $0 < \delta_1 \leq 1$.*

2.2 Histories, strategies and beliefs

The histories after the i^{th} period are $H_i := (A \times R)^i$. The whole space of histories up to the last period is the disjoint union $H := \bigsqcup_0^{k-1} H_i$.

The respondent observes past play, so his strategy in period i is a function of H_{i-1}. His global strategy is a function of the space of histories H. Take this to be a behaviour strategy, giving a mixed action in $\Delta(R)$ at every history: his stragegy is a function $s_2 : H \to \Delta(R)$ such that $s_2(.)(\sigma)$ is a measurable function $H \to R$ for any measurable $\sigma \subseteq R$. Throughout this paper $\Delta(X)$ for any measure space X denotes the space of probability measures on X.

The signaller, in addition to observing past play, knows his current and previous period-types. So define: $HT_i := (A \times R)^i \times T^{i+1}$; $HT = \bigsqcup_0^{k-1} HT_i$. ("Histories with types".) His strategy is a function $s_1 : HT \to \Delta(A)$ satisfying: $s_1(.)(\alpha)$ is a measurable function $HT \to R$ for any measurable $\alpha \subseteq A$.

Since we are dealing with continuous action spaces for both players the measurability assumptions above are needed to be able to define the progress of the game given the strategies. See the section below on the outcome of the game and the corresponding definitions in the appendix to see why this is so.

The respondent at any history $h_i \in H_i$, $i < k$, has a belief $\beta(h_i) \in \Delta(T^{i+1})$ about the signaller's types up to that point.

There is an exogenously given distribution of types in which there is correlation between types from one period to the next. This process will be assumed to be Markov. This correlation will give the motive for the signaller to signal a higher type in a given period: by signalling a higher type he will be thought to be a higher type in the next period. A special case is when types are constant across periods. But we are particularly interested in processes which have full support, so that given any type distribution in a given period, the type distribution in the next period has full support. The equilibrium refinement that I will propose solves the game under this assumption.

The regeneration process is described by the function $\Psi : \sqcup_i T^i \to \Delta(T)$. If types from periods 1 to i are described by $t^i \in T^i$, $\Psi(t^i)$ describes the distribution of types in period $i+1$.

Assumption 9 *Monotonic Markovian type-change:*

$\Psi(.)$ is a Markov process, generated by the function $\psi : T \to \Delta T$ and the initial distribution $\Psi(())$.

$\psi(t)$ is strictly increasing in t in the sense of first-order stochastic dominance.

Assumption 10 *Type regeneration: $\Psi(t^i)$ has full support for any $t^i \in T^i$.*

2.3 The outcome of the game

Define an outcome of the game to be a vector of actions of each player and period-types of the signaller, i.e. an element of $O = (T \times A \times R)^k$, describing the entire progress of the game. Once we have strategies s_1, s_2 and the type regeneration function Ψ we can define from any point in the game ht_i the probability distribution of subsequent play and the probability distribution of outcomes. Call the first distribution the continuation of the game $C^+(s_1, s_2)(ht_i) \in \Delta(T^i \times A^i \times R^i)$ and the second distribution the completion of the game $C(s_1, s_2)(ht_i) \in \Delta\left(T^k \times A^k \times R^k\right)$.

Even though these are familiar notions in game theory, more care than usual needs to be taken since we are dealing with continuous action spaces. Formal definitions of C and C^+ are given in the appendix. The measurability conditions on s_1 and s_2 are needed here for C and C^+ to be defined.

2.4 Perfect-Bayesian Nash equilibrium

We shall be examining perfect-Bayesian Nash equilibria of the game described above. Although the concept is standard, I give a formal definition which respects the particular construction of this model.

Definition 3 (s_1, s_2, β) *is a perfect-Bayesian Nash equilibrium if:*

1. For each history with types ht_i, s_1 maximizes $\int U_1 d[C(s_1, s_2)(ht_i)]$

2. At any history h_i, s_2 maximizes:

$\int\int U_2 d[C(s_1, s_2)(h_i, t^{i+1})]d[\beta(h_i)(t^{i+1})]$

3. For any $h_i \in H_i$, $\beta(h_i)$ satisfies:

3.a) $\beta(h_i)$ respects Bayes' rule where applicable:

If $\beta(h_i)$ gives positive probability to some vector $t^{i+1} \in T^{i+1}$ of period-types in the first $i+1$ periods, and $s_1(h_i, t^{i+1})$ gives positive probability to action a, then for any r, $\beta(h_i, (a, r))$ is the usual Bayesian update of $\beta(h_i)$.

3.b)$\beta(h_i)$ respects Ψ between period i and period $i+1$:

$\beta(h_i)(\{t^{i+1}\}) = \beta(h_i)(\{t^i\} \times T).\Psi(t^i)(t_{i+1})$

The integral in 1. is expected utility for the signaller. The integral in 2. is the expected utility of the respondent given that (h_i, t^{i+1}) is reached, integrated over beliefs about t^{i+1}.

Types with the same period-types up to period i but different future period-types behave in the same way up to period i. Condition 3.b) means that even when a zero-probability event is observed by the signaller, he should not doubt the regeneration process Ψ but given his assessment of the period-types in periods 1 to i his assessment of future period-types will be consistent with Ψ.

3 Equilibrium selection

3.1 The refinement D_ω

The proposed refinement is based on this motivation: out of equilibrium actions that are sub-optimal for all types are considered to be mistakes made by a type in his perception of the response to those actions.[6] The type considered to be making the mistake a is

[6] It is also possible to define a refinement based on utility loss, as follows: if a type t_n is assigned positive probability at the beginning of a period and t'_n is some other type and if t'_n would lose more utility from taking the action a than t_n then t'_n is assigned probability 0 after a is observed. Propositions

thought to become over-confident about the respondent's immediate response to a and that leads him to do a rather that his optimal equilibrium action. The set of beliefs about responses to a that would cause type t to play a are called the (strictly/weakly) justifying beliefs of a for t. Larger mistakes are considered infinitely more improbable than smaller mistakes, uniformly across types, leading to a type who would have had to make a large error in his perception being considered infinitely less likely than a type who would have had to make a smaller error in order to take action a. The probability of errors is itself infinitely small, and no types make errors in equilibrium. This informal argument supports the following refinement D_ω: if the set of weakly justifying beliefs for type t^* is contained in the set of strictly justifying beliefs for type t^{**}, and t^{**} was considered possible (assigned probability > 0) before a was observed, then t^* is given posterior probability 0. The condition implies that beliefs have to move further from the correct equilibrium beliefs for action a to be justified for t^* than for t^{**}.

3.1.1 Justifying Beliefs

Given a perfect-Bayesian Nash equilibrium (s_1, s_2, β), define at a history h_n the justifying beliefs of an action a for a player with $t^{n+1} \in T^{n+1}$ as follows:

Let $u^* = \int U_1 d[C(s_1, s_2)(h_n, t^{n+1})]$ be the expected utility of the optimal strategy s_1 by the signaller.

Given an action a at history h_n, the strategy of the respondent in the next period is described by some function $\widetilde{r} : R \to \Delta R$ giving the mixed response after the current actions (a, r). If s_2' is a strategy of the respondent, the corresponding function is the function which takes r to $s_2'(h_n, (a, r))$. The reason we describe the respondent's next-period strategy as a function is that he could potentially condition his next-period action on his current action and his current action could be mixed. $\widetilde{r}(.)(\pi)$ will be measurable for measurable π. Now given such an $\widetilde{r} : R \to \Delta R$, define the strategy $s_2'(\widetilde{r})$ as follows: $s_2'(\widetilde{r}) = s_2$ except at $(h_n, (a, r))$ for all r, and $s_2'(\widetilde{r})(h_n, (a, r)) = \widetilde{r}(r)$. This strategy by the respondent is the same as the original one but changed in period $n+1$ to respond to $(h_n, (a, r))$ with $\widetilde{r}(r)$ for any r. So it is changed in the next-period response to the action a by the signaller. Now let $\alpha(a)$ be the set of strategies s_1' of the signaller with $s_1'(h_n, t^{n+1}) = a$. Then $u(\widetilde{r}) = \sup_{s_1' \in \alpha(a)} \int U_1 d[C(s_1', s_2'(\widetilde{r}))(h_n, t^{n+1})]$ is the maximum utility of the signaller in response to s_2', conditional on having to play a at the current history.

Definition 4 *For each binary relation* $\rhd \in \{>, \geq, =\}$, $J^{\rhd}(t^{n+1}, h_n, a) :=$

$$\{\widetilde{r} : R \to \Delta R \textit{ with } \widetilde{r}(.)(\pi) \textit{ measurable for measurable } \pi, \textit{ such that } u(\widetilde{r}) \rhd (u^*)\}$$

Call $J^>$ *the strictly justifying beliefs,* $J^{\geq}$ *the weakly justifying beliefs, and* $J^=$ *the barely justifying beliefs.*

1 and 2 below still hold and the structure of all arguments can remain the same: additivite separability makes these two refinements work in a similar way. While this refinement has the advantage of simplicity, D_ω is potentially generalizable to the non-additive case and also does not rely on cardinal utility.

Note that an action a is optimal if the correct belief about the respondent's response given a -the equilibrium strategy - is a barely-justifying belief.

Given h_n call the set $\{a : J^{>}(t^n, h_n, a) \neq \{\}$ for some $t_n\}$ the justifiable actions. These are the actions that are justified for some type by some possibly erroneous belief about the respondent's response. Some actions may not be justified by any belief, and the respondent's beliefs when confronted with these actions will not be specified by the D_ω criterion below.

3.1.2 Definition of D_ω

Definition 5 *A perfect Bayesian equilibrium (s_1, s_2, β) satisfies D_ω if:*

For any history $h_n = (h_{n-1}, a_n, r_n)$ and types t_1^n, t_2^n,

if $\beta(h_{n-1})$ assigns positive probability to t_2^n,

and if $J^{\leq}(t_1^n, h_{n-1}, a_n) \subseteq J^{<}(t_2^n, h_{n-1}, a_n) \neq \emptyset$,

then $\beta(h_n)$ assigns probability 0 to t_1^n.

4 The Iterated Riley solution

4.1 The Riley map

Consider the standard one-shot monotonic signalling game in which the signaller moves first. Imagine that the signaller has utility $u^E(t, a, r) = v_a(t, a) + \delta_1 v_r(r)$, and the respondent response to the signaller's perceived type is given by a stricly increasing function $r'' : T \to R$. The Riley equilibrium of this signalling game is the perfect Bayesian equilibrium in which types separate minimally. Separation implies that the lowest type τ_0 must take his myopic optimal action $a^*(\tau_0)$. Each subsequent type takes his myopic optimal action, subject to separating from lower types. Given the monotonicity assumptions, it is sufficient to require each type to separate from the previous type. Define $RILEY(r'')$ to be this equilibrium, specifying an action for each type. Given any strictly increasing function $r'' : T \to R$, $RILEY(r'')$ is defined inductively as follows:

Definition 6 $RILEY(r'') : T \to A$

$RILEY(r'')(\tau_0) := a^*(\tau_0)$

$RILEY(r'')(\tau_i) := \arg\max_{a \in B_i} v_a(t_i, a)$, *where*

$B_i = \left\{a \in A : u^E(\tau_{i-1}, RILEY(r'')(\tau_{i-1}), r''(\tau_{i-1})) \geq u^E(\tau_{i-1}, a, r''(\tau_i))\right\}$

Assumption 5 ("undesirable $a_{\max}$") guarantees that the set of actions B_i for which a lower type would not want to pretend to be the current type is non-empty. Existence and uniqueness of the arg max above is guaranteed by single crossing and strict quasi-concavity of u^E. (As we saw earlier, u^E must satisfy Assumptions 4-8 since $u_1 = v_a(t,a) + v_r(r)$ does.)

Each type's action is strictly higher than the previous type's by monotonicity (single crossing and preference for higher responses). So by single crossing, if each lower type does not strictly prefer take the subsequent type's action, then any lower type does not strictly prefer to take the action of a higher type. This is why in the function $RILEY$ above it was sufficient to require each type to separate himself from the previous type.

It turns out that the repeated signalling game can be solved by repeated use of the $RILEY$ function. If the signaller after his i^{th} move is thought to be type t, the respondent's action in period $i+1$ will be a function $r''(t)$ of this t. If we only look at actions of the signaller in period i and of the respondent in period $i+1$ we have utility for the signaller given by u^E; this explains the use of the modified utility function u^E in the definition above. Then given the response function r'', the signaller will take Riley separating equilibrium actions $RILEY(r'')$ in period i.

Note that the Riley equilibrium is defined without reference to any distribution of types of the signaller. This fact is very important for analysis of the repeated game and generates history-independence for the signaller.

4.2 The Iterated Riley solution

Under assumptions 1-9 we can now define `the` "Iterated Riley equilibrium", a description of play of both players on the equilibrium path. Assumption 10 (full support) will be used later on to justify the Iterated Riley solution uniquely; it is not necessary to define it. In the Iterated Riley solution the signaller's strategy is a function only of his current type and the stage of the game. His action is given by $\sigma_1 : \{0, ...k-1\} \to T \to A$. $\sigma_1(j)(t_{k-j})$ will define the action of type t_{k-j} of the signaller in the $(k-j)^{th}$ period.

Let $f(\sigma) : T \to T \times A$, $f(\sigma)(\tau) = (\tau, \sigma(\tau))$, so that if σ is a map from period-types to actions, $f(\sigma)$ gives the type and action pair for any type.

σ_1 is defined inductively as follows:

Definition 7 $\sigma_1(0) = a^*$

Given $\sigma_1(j)$, $\sigma_1(j+1) := RILEY\left(r''_{k-j}\right)$, *where* $r''_{k-j}(\tau) = r^*\left(\psi(\tau) \circ f(\sigma_1(j))^{-1}\right)$.

$f(\sigma_1(j))$ represents the map from types to type-action pairs in period $k-j$. Given type τ was believed to have been the signaller's type in period $k-j-1$, the beliefs about the type in period $k-j$ will be $\psi(\tau)$ and the belief about the type-action pair will be $\psi(\tau) \circ f(\sigma_1(j))^{-1}$. $r''_{k-j}(\tau)$ will be the myopic optimal action of the respondent in period $k-j$, given that the signaller is thought to have been type τ in period $k-j-1$.

Assuming that $\sigma_1(j)$ is a strictly increasing function, r''_{k-i} is a strictly increasing function by assumption 3 (increasing responses to types and actions) and assumption 9 (monotonic Markovian type change) and strictly increasing $\sigma_1(j)$, allowing $RILEY\left(r''_{k-j}\right)$ to be defined, which gives a strictly increasing function $\sigma_1(j+1)$. This justifies the definition.

The Iterated Riley solution can now be defined in terms of σ_1:

Definition 8 *s_1, s_2 are an Iterated Riley equilibrium if for histories on the equilibrium path:*

$s_1\left((a_1,...a_{i-1}),(r_1,...r_{i-1}),(t_1,...t_i)\right) = [\sigma_1(k-i)(t_i)]$*, where* $[a]$ *is the degenerate probability measure placing all weight on* a*.*

$s_2\left((a_1,r_1)...(a_i,r_i)\right) = r^*(\psi(\sigma_1(k-i)^{-1}(a_i)) \circ f(\sigma_1(k-(i+1)))^{-1})$ *for* $i \geq 1$

$s_2(()) = r^*\left(\Psi(()) \circ f(\sigma_1(k-1))^{-1}\right)$

To understand the nature of the Iterated Riley solution, consider first these conditions on s_1, and s_2 given σ_1. The signaller's strategy (on the equilibrium path) is described very simply by σ_1: $\sigma_1(j)$ gives the map from types to actions in the period $k-j$. It is independent of previous play and only dependent on the period and the current period-type. The respondent's strategy has a more involved definition. In period $i+1$, he looks at the signaller's last action, a_i. Since we are on the equilibrium path this will be in the image of $\sigma_1(k-i)$. Since σ_1 is strictly increasing it is injective and so only one type $\sigma_1(k-i)^{-1}(a_i)$ will ever take that action. Beliefs about the period-type in the next period should be[7] given by $\psi(\sigma_1(k-i)^{-1}(a_i))$. Since $\sigma_1(k-(i+1))$ gives the map from types to actions in the current period $i+1$, the expected type-action pair will be $\psi(\sigma_1(k-i)^{-1}(a_i)) \circ f(\sigma_1(k-(i+1))^{-1})$. The respondent's action is the myopic response r^* to this. In the first period, the expected type distribution is $\Psi(())$ and the respondent's action is then $r^*\left(\Psi(()) \circ f(\sigma_1(k-1))^{-1}\right)$.

Now consider the definition of σ_1. In the last period, subject to no signalling motives, the signaller takes the myopic optimal action given by the function $\sigma_1(0) = a^*$. If after period $i = k-j$ the signaller is believed to have period-i-type τ, he can expect the response $r^*\left(\psi(\tau) \circ f(\sigma_1(j))^{-1}\right)$. There is a minimal separating equilibrium looking only at actions in the current period and responses in the next, given by the $RILEY$ map applied to this response function and using utility u^E with the response discounted by the discount factor δ_1.

It is useful to specify a map F that gives σ_{i+1} in terms of σ_i. Let the space of strictly increasing functions from T to A be $Inc(T,A)$.

Definition 9 $F: Inc(T,A) \to Inc(T,A)$*,*

$F(\sigma) := RILEY\left(r''\right)$*, where* $r''(\tau) = r^*\left(\psi(\tau) \circ f(\sigma)^{-1}\right)$*.*

[7] I have not specified beliefs in the Iterated Riley solution, and they are mentioned here as an aid to understand the definition.

Then we have $\sigma_1(i) = F^i a^*$.

Note that the correspondence between the signaller's type and his action in the period j periods from the end is the same across games with a varying number of periods, all other specifications constant.

5 The Iterated Riley equilibrium and the D_ω condition

Here I will show the existence and uniqueness of the Iterated Riley equilibrium as a perfect Bayesian-Nash equilibrium satisfying D_ω.

5.1 Supportability of the Iterated Riley solution

Proposition 1 *Under assumptions 1 to 9, iterated Riley solution is supportable as a perfect-Bayesian Nash equilibrium satisfying D_ω.*

Proof. See appendix. ■

Note that the Assumption 10 (full support) is not necessary to support the Iterated Riley solution as a perfect-Bayesian Nash equilibrium satisfying D_ω.

A particular Bayesian-Nash equilibrium is defined explicitly in the proof which satisfies the required properties. It has these properties:

At any point in the game (not only on the equilibrium path but at all histories with types) the signaller takes an action given by σ_1. At a history h_i in which the signaller has taken actions $(a_1, ...a_i)$: respondent's belief about the signaller's period-i type is $[\tau_j]$ if $a_i = \sigma_1(k-i)(\tau_j)$. The respondent's belief in any period is a monotonic function of the previous action, and is always supported on a single type. Beliefs about the period-type t_i are unchanged after period i, on and off the equilibrium path. Beliefs about the period types after period i at history h_i are deduced from the Markov process ψ. See the section below on uniqueness for an explanation of why these beliefs satisfy D_ω. Given these beliefs, the respondent then acts myopically based on his beliefs about the type and action he can expect in the current period. This is because the action that he takes has will have no effect on the future course of the game.

Now suppose that separation is from previous types is always binding. Then beliefs have a particularly simple form: if a_i lies in $[\sigma_1(k-i)(\tau_j), \sigma_1(k-i)(\tau_{j+1}))$ beliefs about the type are still $[\tau_j]$: the respondent assumes it is the lower type making a mistake and taking too high an action rather than a higher type taking too low an action. Below $\sigma_1(k-i)(\tau_0)$ the type is believed to be $[\tau_0]$ and above $\sigma_1(k-i)(\tau_h)$ the type is believed to be $[\tau_h]$.[8]

There will be other equilibria than the one checked that satisfy D_ω. But D_ω does specify the equilibrium up to responses to unjustifiable actions. The signaller's strategy must be given by σ_1 for a D_ω equilibrium. The respondent must respond and form beliefs as above after a *justifiable* actions; after an (out-of-equilibrium) unjustifiable action he may form any beliefs and act accordingly.

5.2 Uniqueness

Proposition 2 *Under Assumptions 1 to 10, in a perfect-Bayesian Nash equilibrium of the model described in section 1 satisfying D_ω: The signaller's strategy depends only on the period and his type in that period via the function σ_1 defined above, by the equation* $s_1\left(\left(a_1, \ldots a_{i-1}\right), \left(r_1, \ldots r_{i-1}\right), \left(t_1, \ldots t_i\right)\right) = \left[\sigma_1(k-i)\left(t_i\right)\right]$.

The respondent's strategy satisfies $s_2\left(\left(a_1, \ldots a_i\right), \left(r_1, \ldots r_i\right)\right) = r^*(\psi(\sigma_1(k-i)^{-1}\left(a_i\right)) \circ f(\sigma_1(k-(i+1)))^{-1})$ *whenever* $a_i \in \operatorname{Im}(\sigma_1(k-i)^{-1})$.

And $s_2(()) = r^*\left(\Psi(()) \circ f(\sigma_1(k-1))^{-1}\right)$.

Proof. See Appendix. ■

Note that this equation for the signaller now holds at every history, not only on the equilibrium path. The proposition implies an equilibrium satisfying D_ω must be an Iterated Riley equilibrium.

Two facts about D1 equilibria should be called to mind to understand how the Iterated Riley solution is selected by the D_ω criterion. Firstly as discussed earlier the Riley equilibrium selected is independent of the initial type distribution. A second and related fact is that the beliefs of the respondent are categorical and regardless of the initial distribution assign probability 1 to some type [9]. The logic of D1 is strong enough to outweigh any disparities in the probabilities of initial types: to express this in terms of the intuitive understanding given above of the divinity criterion, a larger mistake is infinitely less likely uniformly across types than a smaller one, so if one type would require a larger mistake to justify an observed action than another, then the latter is considered infinitely more probable, and so the first type is given probability 0 regardless of how much more likely he was than the second type before the action was observed.

This same logic applies for D_ω in the repeated game. We will have at every stage a single-stage signalling game and regardless of the history at any particular stage - regardless of the current type-distribution ascribed to the signaller by the respondent - there will be the same map from types to action given by the Riley equilibrium. And

[8] Cho, Sobel [6] claim that beliefs of this form always generate a D1 equilibrium of the single-stage monotonic signalling game. This is not quite true: it is only true when separation from previous types is a binding constraint in the Riley equilibrium.

[9] For all important actions of the signaller: the "justifiable" ones in my terminology.

while previous action by the signaller will alter the type-distribution expected in a given period, beliefs by the respondent after the current action will be a function of that action only and will be categorical in nature, ascribing probability 1 to a particular type.

The game is solved from the last period and the above logic applied at every stage. Each action by the signaller in period i is paired with the respondent's action in the next period. In the last period there is no signalling incentive, and the signaller takes his myopic optimal action a^*. In period i for the signaller and period $i+1$ for the respondent, given that the game has been solved for the remainder of the game (periods $i+1$ on for the signaller and periods $i+2$ on for the respondent) and generated the Iterated Riley solution there, we can analyze the action in period i and response in period $i+1$ in isolation. This game will be monotonic because the respondent rewards the signaller for signalling a higher type (see the definition of the Iterated Riley solution). The analysis of this restricted game is like the analysis of the one-stage signalling game under D1. Separation comes from the fact that if two players were to pool in equilibrium, by taking slightly higher actions each could discontinuously increase the beliefs about him to beliefs whose support has a minimum of at least the higher type. And minimal separation comes from the fact that if a type were to take an action in equilibrium that is not his myopic optimum given that he has to separate, then by moving to this myopic optimum conditional on separation, he will (at least) maintain beliefs about him, and increase his current period payoff.

6 Limit properties of the Iterated Riley equilibrium

First it is useful to note the continuity of the solution with respect to the various primitives defining it.

Fact 3 *The Iterated Riley solution $\sigma_1(i) : T \to A$, for each i, is continuous as a function of v_a, v_r, δ, ψ, r^*.*

To see this, observe that the function $RILEY(r'') : T \to A$ is continuous as a function of u^E and r''. $F : Inc(T, A) \to Inc(T, A)$ is then continuous when considered as a function of v_a, v_r, δ, ψ. a^*. So $\sigma_1(i) = F^i a^*$ is continuous as a function of v_a, v_r, δ, ψ, r^*.

Now let ψ_0 be the degenerate type regeneration function, with $\psi_0(t) = [t]$. Note that the full support assumption was not used in the definition of the iterated Riley solution. We have seen that if we have ψ tend to the degenerate function ψ_0 in which types remain the same with probability 1, the iterated Riley solution will tend to the Iterated Riley solution with $\psi = \psi_0$. Let us now consider the properties of the iterated Riley solution with regeneration function $\psi = \psi_0$, as the number of periods from the end i tends to infinity, for fixed v_a, v_r, δ, r^*.

First it is useful to define "discounted Stackelberg" utility. The undiscounted utility of the signaller with period-type τ in any period if he can and does commit to the action

a and is known to be type τ is: $v_a(\tau, a) + v_r(r^*([(\tau, a)]))$. We can call this Stackelberg utility. If the discount factor of the signaller is not 1 it will be more useful to consider the "discounted Stackelberg" utility: $v_S(\delta_1)(\tau, a) := v_a(\tau, a) + \delta_1 v_r(r^*([(\tau, a)]))$. This is the utility for type τ of the action a taken in the current period and plus the discounted utility of the best response in the next period to the type-action pair (τ, a).

Given δ_1 and τ, call the maximum value of this the discounted-Stackelberg payoff (which exists by continuity of all functions involved), and the unique a that maximizes the expression (unique by the concavity assumption on $v_r^*(a)$) the discounted-Stackelberg action $a_S(\delta_1)(\tau)$.

Assumption 11 *$v_S(\delta_1)(\tau, a)$ is strictly quasi-concave in a for each $\tau \in T$.*

This assumption that discounted Stackelberg utility is strictly quasi-concave is important to the limit analysis. It is satisfied for example when v_a and v_r are strictly concave and $r^*([(\tau, a)])$ is linear in a for each τ. In a work incentives example where a is work and $r^*([(\tau, a)])$ is market wage this would be a natural specification.

Proposition 3 *$\sigma_1(i)$ tends to a limit Σ_1 as $i \to \infty$.*

Σ_1 is characterized as follows:

1. $\Sigma_1(\tau_0) = a^*(\tau_0)$

2. Let h be the highest solution for x of:

$v_a(\tau_j, x) + \delta_1 v_r(r^*([(\tau_{j+1}, x)])) = v_S(\delta_1)\,(\tau_j, \Sigma_1(\tau_j))$.

(There are one or two solutions.)

$\Sigma_1(\tau_{j+1}) = \max(h, a^*(\tau_{j+1}))$

Proof. See Appendix ■

The proof involves inductive application of a dynamical systems argument. I will explain here some features of the process generating σ_1. Suppose that for a particular type τ_j, $\sigma_1(i)(\tau_j)$ tends to a limit $\Sigma_1(\tau_j)$ as $i \to \infty$. If $\sigma_1(i)(\tau_{j+1})$ also tends to a limit $\Sigma_1(\tau_j)$ it must satisfy $\Sigma_1(\tau_{j+1}) = \max(x, a^*(\tau_{j+1}))$, for some x for which $v_a(\tau_j, \Sigma_1(\tau_j)) + \delta v_r(r^*([(\tau_j, \Sigma_1(\tau_j))])) = v_a(\tau_j, x) + \delta v_r(r^*([(\tau_{j+1}, x)]))$.

This is because $\sigma_1(\tau_{j+1})$ is either eventually given by a binding constraint of separation from the previous type, or by the myopic optimum $a^*(\tau_{j+1})$. The first is the "normal" case; the second is a failure of signalling to have any effect due to types that are too far apart.

Now consider the equation in x. The first part $v_a(\tau_j, \Sigma_1(\tau_j)) + \delta v_r(r^*([(\tau_j, \Sigma_1(\tau_j))]))$ is the converged period-utility of type τ_j. The second part $v_a(\tau_j, \Sigma_1(\tau_{j+1})) + \delta v_r(r^*([(\tau_{j+1}, \Sigma_1(\tau_{j+1}))]))$ is the converged utility of pretending to be

$\Sigma_1(\tau_{j+1})$. Assuming that the need to separate is a binding constraint, these two must be equal. But even if we know that $\sigma_1(i)(\tau_{j+1})$ converges we have not found what it converges to yet because this equation may have more than one solution: it may have one or two solutions. The lower solution lies below $\Sigma_1(\tau_j)$ if the respondent has a direct preference for rewarding higher types, and so in this case we can rule it out because the limit map from types to actions must be weakly increasing. But in the reputation case where the respondent does not care directly about the signaller's type it is not so easy to rule out the lower solution. It may be the case that both $\Sigma_1(\tau_j)$ and some higher action are solutions to the equation above. For a description of how this is resolved and the higher solution is chosen, see the section on reputation below.

Note that the convergence is not monotonic: this has been confirmed by numerical computation of an example.

7 Reputation

Now make the assumption that the respondent does not care directly about the signaller's type, only about his action:

Assumption 12 *$u_2(t, a, r)$ is a function of a and r only*

It follows that $r^*(\mu_{ta})$ only depends on the probability distribution over actions. Define $v_r^*(a) := v_r(r^*(\mu_t * [a]))$, which is independent of μ_t.

Let the highest action that gives the same Stackelberg utility for type τ as action a be $\overline{a}_S(\delta_1)(\tau, a)$.

Corrolary 1 *$\sigma_1(i)$ tends to a limit Σ_1 as $i \to \infty$.*

Σ_1 is characterized as follows:

1. *$\Sigma_1(\tau_0) = a^*(\tau_0)$*
2. *For each j, $\Sigma_1(\tau_{j+1}) = \max(\overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j)), a^*(\tau_{j+1}))$*

This follows simply from proposition 3, noting that $r^*([(t, x)])$ is independent of t and so that h in proposition 3 is equal to $\overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j))$. In words, if $\Sigma_1(\tau_j)$ is weakly above the discounted-Stackelberg action of type τ_j, then $\Sigma_1(\tau_{j+1}) = \Sigma_1(\tau_j)$, assuming this is above the myopic-optimal action $a^*(\tau_{j+1})$. Otherwise $\Sigma_1(\tau_{j+1})$ jumps up above the discounted-Stackelberg action of type τ_j to $\overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j))$.

Assume that separation is binding (as I show in the proof of proposition 4, this will be true if types are close together). Let us continue the discussion of proposition 3 and examine why if $\Sigma_1(\tau_j)$ is below the discounted-Stackelberg action, $\Sigma_1(\tau_{j+1})$ is equal to $\overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j))$ and not the lower $\Sigma_1(\tau_j)$. Both are solutions of the equation

$v_a(\tau_j, \Sigma_1(\tau_j)) + \delta v_r(r^*([(., \Sigma_1(\tau_j))])) = v_a(\tau_j, \Sigma_1(\tau_{j+1})) + \delta v_r(r^*([(., \Sigma_1(\tau_{j+1}))]))$. There are two facts that combine to give this result. Firstly at every stage $\sigma_1(i)$ is strictly monotonic: there is separation of types. (A separation that does not always occur in the limit as we have seen.) Secondly $\Sigma_1(\tau_j)$ is below the Stackelberg action of type τ_j (this is the case we are considering): it follows that for type τ_j small increases in expectations about his action are more valuable than small increases in his action are painful, and so in order to be thought to be taking the action $\Sigma_1(\tau_j) + x$ for small x, type τ_j would be willing to increase his action to $\Sigma_1(\tau_j) + y$ where $y \geq x$. This is the logic that generates type τ_{j+1}'s action: what would type τ_j be willing to do in order to be thought of as taking type τ_{j+1}'s action. So if $\sigma_1(i)(\tau_j)$ is close to $\Sigma_1(\tau_j)$ and $\sigma_1(i)(\tau_{j+1}) - \sigma_1(i)(\tau_j)$ is small, then $\sigma_1(i+1)(\tau_{j+1}) - \sigma_1(i+1)(\tau_j)$ must be larger. The first difference is the increase in expected action that type τ_j will gain in pretending to be type τ_{j+1}; the second is the increase in action that is necessary. This means that $\sigma_1(i)(\tau_j)$ can never become close to $\sigma_1(i)(\tau_{j+1})$ and is pushed away from $\Sigma_1(\tau_j)$.

Now we can see what happens when types become dense: the main reputation result is for this case. Suppose that u_1 and u_2 are defined continuously over an interval $\overline{T} = [\tau_{\min}, \tau_{\max}]$ and satisfy the relevant assumptions above. Define the function S as follows:

Definition 10 $S : \overline{T} \to A$

$$S(\delta_1)(\tau) = \max(\overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min})), a_S(\delta_1)(\tau))$$

I.e. $S(\delta_1)(\tau)$ gives maximum discounted Stackelberg utility to type τ over the set of actions that give type $\tau_{\min}$ at most the discounted Stackelberg utility of $a^*(\tau_{\min})$. See the diagram Fig 1. Proposition 4 asserts that as types become dense, the limit map from actions to types $\Sigma_1$will converge to $S(\delta_1)$, apart from the lowest type who must take the action $a^*(\tau_0)$.

Proposition 4 *Given $\eta > 0$, we can find $\epsilon > 0$ such that for any finite set $T \subset \overline{T}$ with* $\max_{\overline{\tau}} \min_{\tau}(\{|\tau - \overline{\tau}|, \tau \in T, \overline{\tau} \in \overline{T}\}) < \epsilon$, *the limit solution Σ_1 satisfies*

$$|\Sigma_1(\tau) - S(\delta_1)(\tau)| < \eta \text{ for } min(T) \neq \tau \in T.$$

Proof. See appendix ■

My result is that as the number of periods from the end tends to infinity the given limit holds. The model studied is finitely repeated and over games with different numbers of periods but the same specifications otherwise play is determined by the number of periods from the end. An implication is that for any levels of patience, as the number of periods $k \to \infty$, play in period p converges to the limit found. If the number of periods is large, reputation will take a long time to die out.[10]

[10] "Reputation effects" actually are strong up to the penultimate period. However the given limit properties are only realized further back in the game.

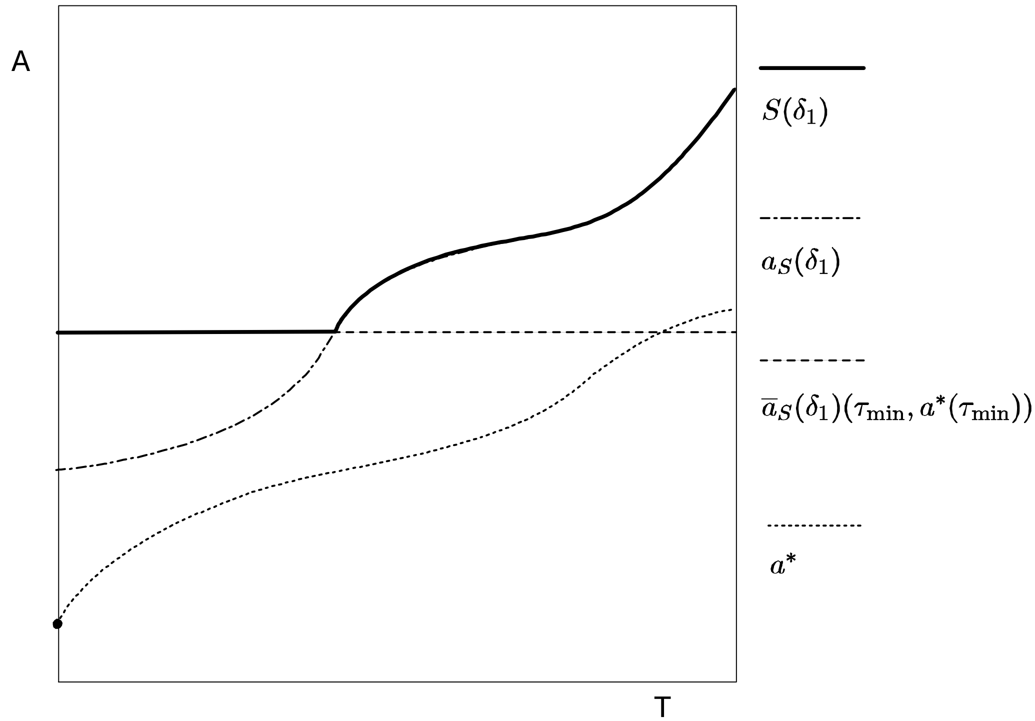

Figure 1: the limit map $S(\delta_1)$

7.1 Discussion

7.1.1 A modified Stackelberg property

The reputation result above combines a separation property for low types with a (discounted) Stackelberg property and logic. The separation property is that types above the lowest type must separate from the lowest type making him unwilling to move from his myopic optimal action $a^*(\tau_{\min})$ and pretend to be a high type, where this willingness is evaluated with (discounted) Stackelberg utility. And types whose (discounted) Stackelberg actions lie above this point take these actions. One can think of the actions that are in $Im(\Sigma)$ - actions that are taken in the limit - as the actions that the signaller can commit to: by taking an action that is in (or more exactly close to something in) this set far from the end of the game, he we be expected to take (close to) the same action in the next period. Thought of in this way the reputation result is that the limit Σ exists and $Im(\Sigma)$ becomes dense in $[\overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}), a_S(\delta_1)(\tau_h)]$ but has a gap in $(a^*(\tau_{\min}), \overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))$ and these actions no-one can commit to. This results in the lower types pooling at $\overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))$.

This result is distinguished from standard reputation models in that these models will just generate a Stackelberg property for the normal type or types: the actions that a normal type can effectively commit to are those actions which get played by behavioural types and these tend to be assumed to include the Stackelberg action of the normal type.

Depending on the context both parts of the curve $S(\delta_1)$ may be interesting, or only one of the two, pooling or Stackelberg. If we think, following a line of thought that is found in the standard reputation literature, that some of our types are "normal" (probable) and others improbable, and that we can imagine a type that is so low that he would rather take his myopically optimal action than commit to a Stackelberg action of a normal type, and we give this low type some positive bur low probability, then our "normal" types will take their Stackelberg actions.

In the lower part of the curve $S(\delta_1)$, types pool at a point determined by separation from the lowest type. This action is higher than the actions that they would like to commit to. But by taking an action even slightly lower than this action, they pay a heavy cost: they are considered to be the lowest type. If they take at least the action $\alpha = \overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))$, they will be expected to take at least this action in the next period, while if they take an action less than α they will be expected to do $a^*(\tau_{\min})$ in the next period. This is a discontinuity that makes taking at least the action α very important. It is appropriate to call α a *reputational standard*, a mark that it is important to reach in order to prevent one's reputation from being destroyed altogether - at least for the next period which is as long as reputations last in this model.

This *reputational standard*, a novel consequence of the repeated signalling model, can potentially be used to explain various situations in which there is a standard of behaviour that can be thought of as a standard necessary to live up to in order maintain a reputation. For example obedience to some social (legal or moral) or business norms can be understood as a requirement for establishing that one is not a bad (criminal or untrustworthy or undependable) type. The point is that these norms are often not continuous but discrete: either one complies with them or one does not.

Let me offer some potential examples in more detail. High actions by the signaller could represent good behaviour by and the respondent could be society; low actions of society could be imprisonment for protection of society, high actions the ability to participate fully in society. Low types are criminal; high types are upstanding citizens. The lowest type will commit crimes; higher types follow the "norm" of the society, which is the particular standard of behaviour, determined endogenously α by the need to separate from the most criminal type. Many people will follow this norm and do no more than this: they follow it because they do not want to be considered the criminal type. And there may be high types that do more both out of natural inclination and the rewards of being thought of as especially trustworthy and good to deal with.[11] Business norms may be thought of in the same way, with the low type being the laziest or least trustworthy.

[11] In contracts, Socrates, justice is of use. - *Plato, The Republic*

It may be more realistic to assume that in addition to treating a person favourably for being expected to behave well, society will be well disposed to a person who is thought to be "good". The repeated signalling model can still deal with this situation. We can still use the same mathematical definition of the *reputational standard* and it will still be the action necessary to avoid being thought of as the worst type. However there will no longer be pooling at this action in the limit.

Relatively lazy types may for example prefer to work shorter hours than 9 to 5, and would be willing to take cuts in pay to do this, but they do not because the reputation cost of not coming up to the fixed standard is steep.

7.1.2 Technical comparison to other models of reputation

The repeated signalling model is fully calculable at all histories of the game. While this is true of some standard reputation models ([17]) it is rare and usually even the actions of the reputation-builder in equilibrium are often not fully specified: limit results tend to be in the form of lower bounds on payoffs of the normal type rather than convergence of his behaviour. By contrast the modified Stackelberg property above specifies the signaller's action in the limit.

The calculability of the model extends to arbitrary discount rates of both players. The respondent's discount factor has no effect on the course of the game. The signaller's discount factor does and the reputation results are given in terms of δ_1. A simple modification of Stackelberg payoffs into "discounted Stackelberg payoffs" is all that is required. This generality is very unusual in the reputation literature, which invariably requires that the patience of the reputation-builder goes to 1. Stackelberg results obtain when the respondent is short-lived ([11]) and there exist limit results in the case when the uninformed player's patience tends to 1 but when the informed player's becomes at the same time infinitely more patient than the uninformed ([10]): $(\delta_2, (1-\delta_1)/(1-\delta_2)) \to (1, 0)$. The special case of strictly conflicting interests ([7]) is an exception, as is the reputational bargaining model of Abreu, Gul [1], in which both discount factors tending to 1 with differences in patience tending to a limit. In general, however, reputation results require the informed player to become infinitely patient and infinitely more patient than the uninformed player. When the opponent is long lived with a fixed discount factor player may be able to establish reputations for complex strategies under certain conditions and do better than the static Stackelberg payoff ([8]). This does happen in the repeated signalling model because reputation is established along one dimension only.

Standard reputational models are often completely general in the stage games studied, while the repeated signalling model analyzes only a class of games in which stage game payoffs are monotonic and additively separable. But within the class I define the model can be completely solved and the question of what happens for any levels of patience of both players addressed, questions which are not addressed in the standard literature. I find that a reputation can be established against a patient player, even by a player that is less patient. And I find that a "discounted Stackelberg" result applies when the informed player is not patient (subject to separation from the lost type). The discounted Stackelberg action is a novelty and just as easy to calculate as the Stackelberg action and can easily be applied to situations in which players are thought of as impatient. One implication of the discounted Stackelberg result for high types is that a small reduction in the discount factor from 1 has a second order effect on (limit) payoffs of the informed player.

The way in which reputation is established is quite different in the repeated signalling model from the standard approach to reputation. Reputation is a one-period property of the repeated signalling model, with the expectation of the respondent being based on the previous action of the signaller, and the signaller can gain or lose it immediately at any time. This happens because it is easier, given the full support assumption, for a type to change than to make a (larger) mistake. For a discussion of signalling without type regeneration, in which this logic does not apply, see section 9.1. The property that reputation can be gained or lost at any point is shared with reputational models with imperfect observability but with much more sudden gains and losses. Without imperfect observability, in standard reputation models, either reputation is lost immediately if at all (revelation of the normal type) or the play from any point in the game tree may be very unknown.

The nature of the types I consider to be an advantage of the repeated signalling model. The commitment types of reputation models are often considered to be an unsatisfactory element, out of place in a theory based on strategy and rationality. On the other hand it has been argued[12] as a genericity assumption they make models involving them at small levels more reasonable than purely "rational" models without. My view is that including behavioural types at small levels is an unobjectionable and valid method, but that the order of the limits involved in reputational models restricts how small the probabilities of behavioural types can be to be effective. The results are found in general under a limit as the informed player becomes infinitely patient for a given probability of behavioural types. If this probability is very small, the required patience may be very large indeed. If we look at the set of discount factors which result in payoffs a certain distance from the Stackelberg (assuming a model that gives a Stackelberg result) as a function of the probability p of behavioural types, we only know that this set contains $(\delta(p), 1)$ for some $\delta(p)$, and this could tend to the empty set as $p \to 0$. For any specified situation with a given high level of patience we will need the probability of behavioural types to be high enough to justify applying a reputation result to expect actions that are near Stackelberg.

The workings of the equilibrium differ from those in reputational models in that the normal type in reputational models pools with commitment types while in the repeated signalling model "normal types" separate from each other in each period.

Mailath and Samuelson [20] they have a model of reputation in which the lowest type's action is fixed and the higher type establishes a reputation by separating himself from the lower type. They find that the higher type will take higher actions than he would otherwise, which may be higher or lower than the Stackelberg action. This holds true in this model with two types: here the lowest type's action is effectively fixed, although he is not a behavioural type, and the higher type separates and may take an action that is more or less than the Stackelberg action depending on the distance between the types. But the most interesting results in the repeated signalling model come from having a large number of types. While with two types separation from the lowest type determines the answer: with more types this model gives both the logic of separation from the lowest

[12] This argument is made I believe by Fudenberg; I will give a reference and exact quotation when I have located the article.

type (for low types) and Stackelberg actions (for high types).

8 The general case

Suppose now that $u_2(t,a,r)$ depends on all three arguments so that $r^*([t,a])$ is a function of both t and a. Consider for example the work-incentives model above. The observable productivity a of the worker (signaller) could be measured by the market (respondent) as quantity of writing, or some other other easy and imperfect measure. Suppose that t is the ability of the worker, with more able workers being better able to produce more writing. It is a reasonable assumption that a worker who produces a given amount of writing has a value to the market that is an increasing function of his ability. The wage $r^*([t,a])$ paid will then be strictly increasing in both t and a.

Assumption 13 *$r^*([(\tau,a)])$ is continuously differentiable in (τ,a) with both partial derivatives strictly positive. $v_r(r)$ and $r^*[(\tau,a)]$ are continuously differentiable. $v_a(\tau,a)$ is differentiable with respect to a, with derivative continuous in (τ,a).*

Definition 11 $G:\overline{T}\to A$

$$G(\delta_1)(\tau_{\min}) := \overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))$$

$$G(\delta_1)'(\tau).[\tfrac{\delta}{\delta a}v_a(\tau,a)+\delta_1\tfrac{\delta}{\delta a}(v_r\circ r^*)([(\tau,a)])]+\delta_1\tfrac{\delta}{\delta \tau}(v_r\circ r^*)([(\tau,a)])$$

The limit will now be given by G instead of S. $G(\delta_1)$ is differentiable and lies above the discounted Stackelberg curve $a_S(\delta_1)$.

Proposition 5 *Given $\eta>0$, we can find $\epsilon>0$ such that for any finite set $T\subset\overline{T}$ with $\max_{\overline{\tau}}\min_\tau(\{|\tau-\overline{\tau}|,\tau\in T,\overline{\tau}\in\overline{T}\})<\epsilon$, the limit solution Σ_1 satisfies*

$|\Sigma_1(\tau)-G(\delta_1)(\tau)|<\eta$ for $min(T)\neq\tau\in T$.

Proof. See Appendix ■

Consider the one-stage signalling game with the same specifications with the signaller moving first and utility given by u^E. In the reputation case we get the discounted Stackelberg action as the solution since the signaller is a Stackelberg leader and type-inference has no significance. In the general case the solution is given by H, say, where H satisfies the differential equation $H'(\tau).[\frac{\delta}{\delta a}v_a(\tau,a)+\frac{\delta}{\delta a}(v_r\circ r^*)([(\tau,a)])]+\frac{\delta}{\delta\tau}(v_r\circ r^*)([(\tau,a)])$ with initial condition $H(\tau_{\min})=a_S(\delta_1)(\tau)$. The differential equation is the same as the differential equation in the limit above but the initial condition is the discounted Stackelberg action of the lowest type rather than the higher $\overline{a}_S(\delta_1)(\tau_{\min},a^*(\tau_{\min}))$.

The standard one-stage signalling game can be thought of as combining commitment and pure type-signalling - commitment trivially because the respondent observes the signaller's action before moving and "pure" type-signalling because given an known action the signaller would still want to be thought as being a higher type. The limit of the repeated signalling game with simultaneous moves also combines commitment and pure type-signalling. The two solutions satisfy the same differential equation but the initial condition (describing types close to the lowest) is higher in the repeated signalling game because the lowest type takes his myopic optimal rather than his discounted Stackelberg action. Thus apart from the lowest type there is more costly signalling in the repeated game with simultaneous actions than in the one-stage game with the signaller moving first.

9 Further work and extensions

9.1 The fixed type case

If there is no type-regeneration what will happen in the repeated signalling game? The Iterated Riley solution can be defined in this case, but is no longer selected uniquely by the equilibrium refinement D_w. Moreover in an Iterated Riley solution for many histories (off the equilibrium path) the respondent will believe that the signaller's type has changed. If we were to require that the respondent's beliefs always assign probability 0 to type-change, then the Iterated Riley solution is ruled out. (Here the difference is the difference between having a type space T^n with types staying the same with probability 1 and having a type space T. The space of Bayesian Nash equilibria is different because the restrictions on beliefs after probability zero events are different.)

The criterion D_w will be weak because it relies on full support at a given stage for its strength: if a type has probability zero before period i and but has a larger set of justifying beliefs for a given action than any other type, D_w does not specify what beliefs will be after that action is observed. Another criterion would do better. One possibility that I have partly analyzed is measuring for a given type and history the error at each period of the types actions, combining these errors into a real-valued total error via some norm, and specifying that the respondent's beliefs about the signaller's type after a given history have support in the set of types with the least total error. I find that if there is a "reasonable" solution in the sense that higher actions generate higher beliefs about the type and a continuity property holds, then all signalling must happen in the first period: that is to say, on the equilibrium path, myopic actions are taken except in the first period. Equilibrium play is then independent of the norm and measure of error above. However the existence and necessity of such an equilibrium have however not yet been shown.

Kaya [15] has a model of repeated signalling in which the signaller has a fixed type. Rather than using any refinement she calculates the "least cost" separating equilibrium, the separating equilibrium where it exists that is most preferred by all types. Kaya finds that when types are ranked in the convexity of their payoff functions, there will be such an equilibrium, and in this equilibrium signalling will be either spread out or all in the first period depending on the direction of the convexity ranking. The signaller moves first in the stage game and so the model does not study reputational issues.

9.2 Multi-dimensional A

If the action set A of the signaller is a product of intervals in $\mathbb{R}^n$ rather than $\mathbb{R}$, it will be necessary to find an assumption on v_a that gives a (uniquely) definable and strictly increasing function $RILEY\,(r'')$ for any strictly increasing $r'' : T \to R$. If this can be found then propositions 1 and 2 go through with no changes. It will then be interesting to see what the limit properties are. Preliminary work indicatest that if the methods can be extended to $\mathbb{R}^n$ there will be a curve $T \to A$ and the current Stackelberg result will apply when restricted to this curve, but that the curve will be defined by a differential equation unrelated to any commitment property. In the one dimensional case it is the necessity of separating from the previous type at every stage that determines the solution and gives the reputation property. In more dimensions there will be a whole range of actions at any stage that just separate from the previous type; the exact specification of the subsequent type's utility on this set, which will give his Riley action, then becomes significant.

9.3 Two-way uncertainty, non-additive utility

Additive utility is a reasonable specification for some situations but not others. Current analysis indicates that what is needed is an assumption that gives a unique equilibrium at each point in time in the actions of both players. When a good assumption is found the game is solveable by backward induction as before, with monotonicity and single crossing being preserved at each stage. The history-independence property will no longer hold.

Two way uncertainty in the repeated signalling model will require the action of each player to be both dependent on his type and to be a reward for a higher expected type or action of the other player. With additive separability for both players the model will involve two-way type-signalling only. A potential model is reciprocation, with types representing something like good-will and both players being made more generous by their own and the other player's revealed good will. Current work studies such a model, although outside of a limited informational setting it turns out there are difficult problems with preserving monotonicity as the game is solved backwards.

Two-way uncertainty without the additivity assumption would open up a large class of models, with the potential to study reputational incentives on both sides. Very interesting two-way reputational models in have been studied in the context of bargaining ([1], [2]). Applications of two-way signalling models include oligopoly and work incentives in teams.

9.4 A compact set of types

Ramey[22] solves the one-stage signalling game using D1 with any compact set of types in $\mathbb{R}$, generalizing Cho, Sobel [6]. The repeated signalling can be defined and extended in the same way and propositions 1 and 2 (existence and uniqueness) will hold. It will involve a more complex notation and an adapted version of the Riley equilibrium. The arguments for the limit results will need a new approach, but it is possible that something in the spirit of the current inductive proof will work. The benefit of compact continuum of types is that the limit should be exactly the function S and we should not need to take a limit as types become dense. And a continuum of types will often be a more natural specification of a given repeated signalling situation, with a continuum being a natural way to model types with more or less of a certain predisposition, and a continuum of potential types being part of the limit reputation result (u_1 and u_2 being defined on the real interval $\overline{T}$ of types). It is unlikely that this generalization will add new insights into repeated signalling and reputation.

10 Appendix

Definition 12 *$P^{HT\to H}: HT \to H$ projects from histories with types to histories in the natural way.*

$P^{HT\to\sqcup T^j}: HT \to \sqcup T^j$ projects from histories with types to vectors of types in the natural way.

$P_j^{T^i}: T^i \to T$ for $i \geq j$ projects to the j^{th} period-type.

Definition 13 *The completion of the game $C(s_1, s_2): HT \to \Delta\left(A^k \times R^k \times T^k\right)$, given strategies of each player, is defined as follows:*

1. *Let $C(s_1, s_2)(ht_i)[\{\tau \times \left(A^{k-j} \times R^{k-j} \times T^{k-(j+1)}\right)]$ over measurable sets τ of $(A^j \times R^j \times T^{j+1})$ be denoted $\Phi_j(\tau)$.*

Φ_j for $j \in \{i, ...k-1\}$ is defined inductively by:

a. *$\Phi_i([(a_1, ...a_i), (r_1, ...r_i), (t_1, ...t_{i+1})]) = 1$. I.e. the existing history happens with probability 1 in the completion.*

b. *If α, β, γ are measurable in A, R, T respectively,*

$$\Phi_{j+1}[\tau\times(\alpha\times\beta\times\gamma)] = \int_{ht_j\in\tau} s_1(ht_j)(\alpha).s_2\left(P^{HT\to H}ht_j\right)(\beta).\Psi\left(P^{HT\to\sqcup T^j}ht_j\right)(\gamma)d\Phi_j(ht_j).$$

This defines a probability measure Φ_{j+1} given Φ_j.

2. *Then over measurable sets τ of $\left(A^{k-1} \times R^{k-1} \times T^k\right)$, and for α, β are measurable in A, R,*

$$C(s_1, s_2)(ht_i)[\tau\times\alpha\times\beta] = \int_{ht_{k-1}\in\tau} s_1(ht_{k-1})(\alpha).s_2\left(P^{(HT\to H)}ht_{k-1}\right)(\beta)d\Phi_{k-1}(ht_{k-1}).$$

This defines the probability measure $C(s_1, s_2)(ht_i)$.

The integrals above exist by the measurability assumptions on s_1 and s_2. They define measures on the appropriate product spaces.[13] 1. defines Φ_j inductively from $j = i$ to $j = k-1$. 2 defines the probability measure $C(s_1, s_2)(ht_i)$ given Φ_{k-1}.

Definition 14 *The continuation play $C^+(s_1, s_2)$ of the game, given strategies of each player, is defined as follows:*

$C^+(s_1, s_2) : HT \to \sqcup\Delta(A^i \times R^i \times T^i)$, with $C^+(s_1, s_2)(ht_j) \in \Delta\left(A^{k-j}, R^{k-j}, T^{k-j}\right)$

$C^+(s_1, s_2)(ht_i) := C(s_1, s_2)(ht_i) \circ P^{-1}$, where P projects $(A \times R \times T)^k$ onto the last $k-i$ coordinates $(A \times R \times T)^{k-i}$.

Proof. Proof of Proposition 1: Supportability of the iterated Riley solution as a perfect Bayesian equilibrium satisfying $\mathbf{D}_\omega$

Define s_1, s_2, β as follows:

For any $ht_i = ((a_1, ...a_i), (r_1, ...r_i), (t_1, ...t_{i+1}))$, $h_i = ((a_1, ...a_i), (r_1, ...r_i))$:

$s_1(ht_i) = [\sigma_1(k-(i+1))(t_{i+1})]$.

Define $r_i''(\tau) := r^*(\psi(\tau) \circ f(\sigma_1(k-(i+1)))^{-1})$.

Define $\underline{r}_i(\tau, a) := \inf\{r : u^E(\tau, a, r) \geq u^E(\tau, \sigma_1(k-(i+1))(\tau), r_i''(\tau))\}$. (The infimum is taken over the set R so that $\inf(\emptyset) = r_{\max}$.). $\underline{r}_i(\tau, a)$ will be the minimum response in the next period that would justify action a for type τ.

For $j \leq i$, let $\beta_j(h_i) = [\sup(\arg\min_\tau \underline{r}_j(\tau, a_j))]$. This defines beliefs over the type in the first i periods. Beliefs about the full k-period type are generated from these beliefs by ψ.

Note that $\underline{r}_i(\tau, \sigma_1(k-(i+1))(\tau)) = r_i''(\tau)$ trivially, while by the definition of the Riley equilibrium, $\underline{r}_i(\tau', \sigma_1(k-(i+1))(\tau)) \geq r_i''(\tau)$ with strict inequality for $\tau' > \tau$: any higher type would strictly lose on moving to the action of a lower type. So if $a_j = \sigma_1(k-j)(\tau)$ then $\beta_j(h_i) = [\tau]$.

Let $s_2(h_i) = r_i''(\beta_i(h_i))$, player 2's myopic best response response to the action a_i in period n given above beliefs.

Now I will show that (s_1, s_2, β) is a perfect Bayesian equilibrium. It follows that it is an Iterated Riley equilibrium and so the proposition is proved.

[13] For uniqueness of the defined measure, note that we have defined the measure on all product sets, which are a π-system generating the product σ-algebra. For existence, we have defined a measure on product sets. Extend additively to a measure on the class of finite disjoint unions of product sets (uniquely). Applying the monotone convergence theorem, this measure is countably additive on this class, which is a ring of sets generating the product σ-algebra, and so extends to a measure on the product σ-algebra by Caratheodory's extension theorem.

Claim 1 *(s_1, s_2, β) is a perfect Bayesian equilibrium*

Optimality of s_1

Let $\widehat{r}_{j+1}(a_i)$ be player 2's action in period $j+1$ in response to player 1's action a_j in period j: $\widehat{r}_j(a) := r''_j(\sup(\arg\min_\tau \underline{r}_j(\tau, a)))$.

Player 1's utility at history ht_i given player 2's strategy above, when he takes strategy s'_1 is

$K_1 + \delta^{i+1}[v_r(\widehat{r}_{i+1}(a_i)))] + \delta_1^{i+1}[v_a(t_{i+1}, a_{i+1}) + \delta_i v_r(\widehat{r}_{i+2}(a_{i+1}))] + ... + \delta_1^{k-1}[v_a(t_{k-1}, a_{k-1}) + \delta_i v_r(\widehat{r}_k(a_{k-1}))] + \delta_1^k[v_a(t_k, a_k)]$, integrated over $C(s'_1, s_2)(ht_i)$, where K is a constant (utility up to period i) independent of s'_1.

This expression equals $K_2 + \delta_1^{i+1}[v_a(t_{i+1}, a_{i+1}) + \delta_1 v_r(\widehat{r}_{i+2}(a_{i+1}))] + ... + \delta_1^{k-1}[v_a(t_{k-1}, a_{k-1}) + \delta_1 v_r(\widehat{r}_k(a_{k-1}))] + \delta_1^k[v_a(t_k, a_k)]$, where $K_2 = K_1 + \delta^{i+1}[v_r(\widehat{r}_{i+1}(a_i)))]$ is independent of s'_1.

I will show that $a_j = \sigma_1(k-j)(\tau)$ maximizes $[v_a(\tau, a_j) + \delta_1 v_r(\widehat{r}_{j+1}(a_j))]$ for any τ. It follows from this that player 1's strategy maximizes each component $[v_a(t_j, a_j) + \delta_1 v_r(\widehat{r}_{j+1}(a_j))]$ of the expression above since it puts probability 1 on $a_j = \sigma_1(k-j)(t_j)$. And so it maximizes expected utility of player 1 after any history with types ht_i.

By construction of $\sigma_1(k-j)$, $\widehat{r}_{j+1}$ as the Riley equilibrium with utility $v_a + \delta_1 v_r$, $a_j = \sigma_1(k-j)(t_j)$ maximizes $[v_a(t_j, a_j) + \delta_1 v_r(\widehat{r}_{j+1}(a_j))]$ over $\mathrm{Im}(\sigma_1(k-j))$. (As shown above, $\widehat{r}_{j+1}(\sigma_1(k-j)(\tau)) = r''_j(\tau)$.) It needs to be shown that $a_j = \sigma_1(k-j)(t_j)$ maximizes the expression over all A.

Suppose that for type τ_m, action $a = \alpha$ gives a higher value of $v_a(\tau_m, a) + \delta_1 v_r(\widehat{r}_{j+1}(a))$ than $a = \sigma_1(k-j)(\tau_m)$. Then this is also true for any type with a lower value of $\underline{r}_i(\tau, a)$. So take without loss of generality $\tau_m = \sup(\arg\min_\tau \underline{r}_j(\tau, \alpha))$. Then $\widehat{r}_i(\alpha) = r''_i(\tau_m)$ given beliefs of player 2 specified above.

We know that $\sigma_1(k-j) \geq a^*$. If $\sigma_1(k-j)(\tau_m) \leq \alpha$ then $a^*(\tau_m) \leq \sigma_1(k-j)(\tau_m) \leq \alpha$ and by quasi-concavity of v_a we must have decreasing v_a above $a^*(\tau_m)$. So $v_a(\sigma_1(k-j)(\tau_m)) \geq v_a(\alpha)$. The actions $\sigma_1(k-j)(\tau_m)$ and α both generate the same response by player 2 and so the former gives a (weakly) higher value of $v_a(\tau_m, .) + \delta_1 v_r(\widehat{r}_{j+1}(.))$. This contradicts our assumption, so we must have $\alpha < \sigma_1(k-j)(\tau_m)$.

Now suppose that $\alpha < \sigma_1(k-j)(\tau_{k-1}) < \sigma_1(k-j)(\tau_k)$. Since τ_k prefers $\sigma_1(k-j)(\tau_k)$ and the corresponding response to $\sigma_1(k-j)(\tau_{k-1})$, and the single crossing condition holds between α and $\sigma_1(k-j)(\tau_{k-1})$, we must have $\underline{r}_j(\tau_{k-1}, \alpha) < \underline{r}_j(\tau_k, \alpha)$ contradicting $\tau_m = \sup(\arg\min_\tau \underline{r}_j(\tau, \alpha))$. So τ_k is the minimal type such that $\alpha < \sigma_1(k-j)(\tau_k)$.

We can rule out that $\sigma_1(k-j)(\tau_k) = a^*(\tau_k)$ because if this were so moving below $a^*(\tau_k)$ to α and being thought of as the same type will hurt type τ_k.

So $\sigma_1(k-j)(\tau_k) > a^*(\tau_k)$, which implies by construction of the Riley equilibrium that type τ_{k-1} must exist (that $\tau_k \neq \tau_0$) and that $\sigma_1(k-j)(\tau_k)$ is optimal for type τ_{k-1} (indifference between own action and the next type's).

So by single crossing, comparing $\sigma_1(k-j)(\tau_k)$ with α, we must have $\underline{r}_j(\tau_{k-1},\alpha) < \underline{r}_j(\tau_k,\alpha)$, which again contradicts the definition of τ_k.

So s_1 must be optimal.

Optimality of s_2

It is sufficient for each action in the support of player 2's strategy at any history to be optimal given the rest of player 2's strategy.

It is clear that player 2's action at any history does not affect any subsequent play either of player 1 or of player 2.

Therefore the myopic best response is optimal.

Consistency of beliefs

1. Bayesian updating: If $t^k = (t_1, ...t_k)$ and $\beta(h_i)(t^k) > 0$, then we must have beliefs assign probability 1 to type $t_1, ...t_i$ and which are generated by ψ afterwards. If some type the action a_{i+1} in period $i+1$ with positive probability then $\sigma_1(k-i-1)(t_{i+1}) = a_{i+1}$ for a unique t_{i+1}. On observing a_{i+1}, player 2 assigns probability 1 to $t_1, ...t_{i+1}$ and beliefs about future types are generated by ψ. This is the Bayesian update on the information that the current period-type is t_{i+1}.

2. $\beta(h_i)$ respects Ψ after period i: by definition.

Therefore (s_1, s_2, β) is a perfect Bayesian equilibrium.

Claim 2 *(s_1, s_2, β) satisfies D_ω.*

Suppose not. Then for some history $h_n = ((a_1, ...a_n), (r_1, ...r_n))$, with history $h_{n-1} = ((a_1, ...a_{n-1}), (r_1, ...r_{n-1}))$ in the previous period, and for some $t^n_* = (t^*_1, ...t^*_n)$ and $t^n_{**} = (t^{**}_1, ...t^{**}_n)$ in T^n, $J^{\leq}(t^n_*, h_{n-1}, a_n) \subseteq J^{<}(t^n_{**}, h_{n-1}, a_n) \neq \emptyset$ and $\beta(h_n)$ assigns non-zero probability to t^n_*.

Consider the justifying beliefs for the action a_n for type $t^n = (t_1, ...t_n) \in \{t^n_*, t^n_{**}\}$ at history h_{n-1}.

Let $u^*(t^n) = \int U_1 d[C(s_1, s_2)(h_{n-1}, t^n)]$ be expected equilibrium utility for player 1 of type t^n at the beginning of period n.

Given $\widetilde{r} : R \to \Delta R$ (satisfying the measurability requirement), let $s'_2(\widetilde{r}) = s_2$ except at $(h_{n-1}, (a_n, r))$ for all r with $s'_2(\widetilde{r})(h_{n-1}, (a_n, r)) \equiv \widetilde{r}(r)$. Let $\alpha(a_n)$ be the set of strategies s'_1 of player 1 with $s'_1(h_{n-1}, t^n) = a_n$. Then let $u(t^n)(\widetilde{r}) = \sup_{s'_1 \in \alpha(a_n)} \int U_1 d[C(s'_1, s'_2(\widetilde{r}))(h_{n-1}, t^n)]$, the maximum utility of player 1 in response to s'_2 conditional on having to play a at the current history.

Let $s'_1 = s_1$ except at (h_{n-1}, t^n) where $s'_1(h_{n-1}, t^n) = a_n$.

Now since $s'_2(\widetilde{r}) = s_2$ from period $n+1$ on, and player 2's action in period n affects player 1's utility additively, s'_1 is optimal within $\alpha(a_n)$ against $s'_2(\widetilde{r})$.

Then $C(s_1, s_2)(h_{n-1}, t^n)$ and $C(s'_1, s'_2(\widetilde{r}))(h_{n-1}, t^n)$ differ only in the period n actions by player 1, which are $\sigma_1(k-n)(t_n)$ and a_n respectively, and the period $n+1$ responses by player 2, which are $r''(t_n)$ and $\widetilde{r}(s_2(h_{n-1}))$ respectively.

So $u(t^n)(\widetilde{r}) - u^*(t^n) = [\delta_1^n v_a(t_n, a_n) + \delta_1^{n+1} v_r(\widetilde{r}(s_2(h)))] - [\delta_1^n v_a(t_n, s_1(h_{n-1}, t^n)) + \delta_1^{n+1} v_r(s_2(h_{n-1}, s_1(h_{n-1}, t^n), s_2(h_{n-1})))]$, regarding $s_1(h_{n-1}, t^n)$ and $s_2(h_{n-1})$ as elements of A and R since they are degenerate probability measures.

Take $\widetilde{r}^* \in J^{=}(t^n_{**}, h_{n-1}, a_n)$, which is possible since $J^{<}(t^n_{**}, h_{n-1}, a_n) \neq \emptyset$.

$[v_a(t_n^{**}, a_n) + \delta_1 v_r(\widetilde{r}^*(s_2(h_{n-1})))] = [v_a(t_n^{**}, s_1(h_{n-1}, t_n^{**})) + \delta_1 v_r(s_2(h_{n-1}, s_1(h_{n-1}, t_n^{**}), s_2(h_{n-1})))]$, extending v_r here to expected utility over probability measures.

So $[v_a(t_n^{**}, a_n) + \delta_1 v_r(\widetilde{r}^*(s_2(h_{n-1})))] = [v_a(t_n^{**}, \sigma_1(k-(i+1))(t_n^{**})) + \delta_1 v_r(r''_i(t_n^{**}))]$

So $u^E(t_n^{**}, a_n, \widetilde{r}^*(s_2(h_{n-1}))) = u^E(t_n^{**}, \sigma_1(k-(i+1))(t_n^{**}), r''_i(t_n^{**}))$.

Let $\bar{r} \in R$ s.t. $v_r(\bar{r}) = v_r(\widetilde{r}^*(s_2(h_{n-1})))$. $\bar{r}$ exists uniquely by continuity and monotonicity of v_r.

Then $u^E(t_n^{**}, a_n, \bar{r}) = u^E(t_n^{**}, \sigma_1(k-(i+1))(t_n^{**}), r''_i(t_n^{**}))$ so $\underline{r}_i(t_n^{**}, a_n) = \bar{r}$.

Since $\beta(h_n)$ assigns non-zero probability to t_*^n, $J^{\leq}(t_*^n, h_{n-1}, a_n) \subseteq J^{<}(t^n_{**}, h_{n-1}, a_n)$ by assumption and so $\underline{r}_i(t_n^*, a_n) \leq \underline{r}_i(t_n^{**}, a_n)$.

So $\underline{r}_i(t_n^*, a_n) \leq \bar{r}$ and

$u^E(t_n^*, a_n, \bar{r}) \leq u^E(t_n^*, \sigma_1(k-(i+1))(t_n^*), r''_i(t_n^*))$

$u^E(t_n^*, a_n, \widetilde{r}^*(s_2(h_{n-1}))) \leq u^E(t_n^*, \sigma_1(k-(i+1))(t_n^*), r''_i(t_n^*))$

So $\widetilde{r}^* \in J^{\leq}(t_*^n, h_{n-1}, a_n)$.

So $\widetilde{r}^*$ is in $J^{\leq}(t_*^n, h_{n-1}, a_n)$ but not in $J^{<}(t^n_{**}, h_{n-1}, a_n)$, contradicting our assumption.

QED ■

Proof. Proof of proposition 2 (Uniqueness)

The inductive step, proposition $P(i)$, is defined as follow:

For all histories $ht_j \in HT^j$ for $j \geq i$, i.e. from the $(i+1)^{\text{th}}$ period on, player 1's strategy is described by $\sigma_1(j)$, and this strategy is optimal against s_2 even if player 2's strategy is altered from the equilibrium one in the $(i+1)^{\text{th}}$ period only.

It is trivial that in the last period player 1's strategy is described by a^*. So $P(1)$ is true.

Assume $P(i)$.

Then for $j \geq i$, player 2's action at history $h_j \in H^j$ in period $j+1$ when player 1's action in period j $\left(P_j^{H_j \to A} h_j\right)$ is in the image of $\sigma_j(k-j)$ is as in section II.2.ii.

This is because player 2 at the beginning of period j gave a positive probability to all of player 1's possible types in that period:

If the sub-history of h_j at the beginning of period j is $h_{j-1}\left(=P^{H_j\to H_{j-1}}h_j\right)$, 2's beliefs after observing h_{j-1} about the type of player 1 in period $j-1$ are $\mu=\beta\left(h_{j-1}\right)\left(P^{T^k}_{j-1}\right)^{-1}$.

By the definition of perfect-Bayesian Nash equilibrium, $\beta\left(h_{j-1}\right)$ respects ψ after period $j-1$, so:

The beliefs about player 1's j-type after observing h_{j-1} are $\beta\left(h_{j-1}\right)\left(P^{T^k}_{j}\right)^{-1}=\int\psi(t)(.)d\mu$, which has full support over T since $\psi(t)$ does for each t.

Since only one type takes each action in $\operatorname{Im}\left(\sigma_j\right)$ (by $P(i)$), player 2 observing such an action a in period j assigns probability 1 to the appropriate type $\left(\sigma_j\right)^{-1}(a)$ (as being player 1's j-type), and forms beliefs ν as in section 4.2 about player 1's type and action in period $j+1$.

Since by the assumption $P(i)$ player 2's actions have no effect on player 1's future actions, player 2 acts myopically in each period after i, and so in period $j+1$ takes the myopic best response $r^*(\nu)$ to player 1's expected type-action pair ν.

Now consider a history ht_{i-1}, with type τ_1 at period i. Suppose that at history with types ht_{i-1} we replaced equilibrium strategy for player 1 by $a\in A$, and at history with types $ht_i\oplus(a,r,t)$ for any r,t replaced equilibrium strategy for 1 by $\alpha(r,t)$ for 2 by $\hat{r}(r)$. Assume α is optimal. Then $\hat{r}$ is a justifying belief for type τ_1 at history t_{i-1} if the utility is now at least as great as it was before.

We can take α to be the strategy $\sigma_i(i)$ because we know this is optimal by $P(i)$. Now the strategy of player 1 after period i is fixed, and player 2's actions after period $i+1$ as a function of player 1's type are fixed, independent of a and $\hat{r}$.

Player 1 would then get a continuation utility as a function of a and $\hat{r}$ given by $v_a\left(\tau_1,a\right)+s_2\left(h_{i-1}\right)\left[v_r\right]+\delta\left(\hat{r}(s_2\left(h_{i-1}\right))\left[v_r\right]\right)+const1=v_a\left(\tau_1,a\right)+\delta\left(\hat{r}(s_2\left(h_{i-1}\right))\left[v_r\right]\right)+const2$.

So the utility just depends on the direct utility of a via $v_a\left(\tau_1,a\right)$ and the expectation of the reward in the next period $\hat{r}(s_2\left(h_{i-1}\right))\left[v_r\right]$ via the value of the reward v_r.

We now have the level of simplicity of the two-period signalling game where player 1 moves first and player 2 responds.

No pooling:

Suppose at history h_{i-1} action $\bar{a}$ is in the support of two types with period $-\,i$ types τ and τ', with $\tau<\tau'$, where τ' is maximal. Observing $\bar{a}$, player 2 forms beliefs β about player 1's i-period type that are strictly less than $[\tau']$, resulting in a response $\bar{r}$ in the next period. The action $\bar{a}$ is weakly justified for types τ , τ' by the belief $\hat{r}'(r)=[\bar{r}]$ in period $i+1$. Consider player 2's best response r'' to the belief $[\tau']$ about 1's period $-\,i$ type. Let $\hat{r}''(r)=[r'']$ in period $i+1$. Since $r''>r'$, $\hat{r}''$ strictly justifies $\bar{a}$ for types τ ,

τ', so for small $\epsilon \in \mathbb{R}^n$, $\hat{r}''$ strictly justifies $a'' = \bar{a} + \epsilon$ for types τ , τ'. We will 2's actual strategy in period i+1 as a response to a'' is going to be at least $\hat{r}''$.

Suppose $\hat{r}'''$ weakly justifies a'' for type τ. Let $\hat{r}'''(s_2(h_{i-1}))[v_r]$ be V''' and $\hat{r}'(s_2(h_{i-1}))[v_r]$ be V'.

Then $v_a(\tau, a'') + \delta V''' \geq v_a(\tau, \bar{a}) + \delta V'$. Then by single-crossing $v_a(\tau', a'') + \delta V''' > v_a(\tau', \bar{a}) + \delta V'$ since $a'' > \bar{a}$ and $\tau' > \tau$. So $\hat{r}'''$ strictly justifies a'' for type τ'.

So type τ is assigned probability 0 by player 2 after observing a'', by criterion D_ω, since any belief that would weakly justify his taking action a'' would strictly justify type τ'.

Also note that if $\hat{r}$ weakly justifies action a'' for any type $\tau^- < \tau'$ then it must strictly justify action a'' for type τ', because type τ's utility in equilibrium is at least his utility on taking action $\bar{a}$. So any type $\tau^- < \tau'$ is also assigned 0 probability. So player 2's belief is supported on $\{t \in T : t \geq \tau\}$, so strictly justifies type τ', so the action $\bar{a}$ could not have been optimal for type τ'.

Minimal separation

If a current-type $t(i)$'s strategy involved taking an action that did not maximize $u^E(t(i), a, r''(t(i)))$ subject to separating from lower types then he could change his action to the action that does maximize this subject to separating from lower types and still being perceived as at least type $t(i)$, so can raise utility. (Expand)

The result then follows for period i, and by induction for all periods. ∎

Proof. Proof of proposition 3 (Convergence of $\sigma_1(i)$)

The proof is by induction on the type number, applying dynamical systems arguments for each type assuming the previous type's action converges.

$\sigma_1(i)(\tau_0) = a^*(\tau_0)$ is constant, so tends to the limit $a^*(\tau_0)$.

Suppose $\sigma_1(i)(\tau_j)$ tends to a limit Σ as $i \to \infty$.

σ is defined by:

1. $\sigma_1(0) = a^*$

2. $\sigma_1(i)(\tau_0) = a^*(\tau_0)$

3. Given $\sigma_1(i+1)(\tau_j)$, $\sigma_1(i)(\tau_j)$, $\sigma_1(i)(\tau_{j+1})$, let h be the solution above $a^*(\tau_j)$ of:

$v_a(\tau_j, \sigma_1(i+1)(\tau_j)) + \delta v_r(r^*([(\tau_j, \sigma_1(i)(\tau_j))])) = v_a(\tau_j, h) + \delta v_r(r^*([(\tau_{j+1}, \sigma_1(i)(\tau_{j+1}))]))$

Then $\sigma_1(i+1)(\tau_{j+1}) = \max(\{h, a^*(\tau_{j+1})\})$.

Write the sequence $\sigma_1(0)(\tau_j), \sigma_1(1)(\tau_j), ...$ as $x_0, x_1, ...$ $x_j \to \Sigma$.

Write the sequence $\sigma_1(0)(\tau_{j+1}), \sigma_1(1)(\tau_{j+1}), ...$ as $y_0, y_1, ...$

Let $A(x) := v_a(\tau_j, x)$ and $B_i(x) := v_a(\tau_j, x_{i+1}) + \delta v_r(r^*([(\tau_j, x_i)])) - \delta v_r(r^*([(\tau_{j+1}, x)]))$. A is defined on $[a^*(\tau_j), a_{\max}]$ and is strictly decreasing and continuous on this set.

Then $y_{i+1} = \max(A^{-1}B_i(y_i), a^*(\tau_{j+1}))$.

Define $B_\infty(x) := v_a(\tau_j, \Sigma) + \delta v_r(r^*([(\tau_j, \Sigma)])) - \delta v_r(r^*([(\tau_{j+1}, x)]))$.

An eventual lower bound on the sequence y_{i+1}.

Consider the function $F_i = A^{-1}B_i$. (So that $y_{i+1} = \max(F_i(y_i), a^*(\tau_{j+1}))$.)

Given α and $i \in \{0, 1, ...\}$, consider the set $S_i(\alpha) = \{x : F_i(x) \geq x + \alpha\} = \{x : B_i(x) \leq A(x + \alpha)\}$.

$S_i(\alpha) = \{x : v_a(\tau_j, x_{i+1}) + \delta v_r(r^*([(\tau_j, x_i)])) \leq \delta v_r(r^*([(\tau_{j+1}, x)])) + v_a(\tau_j, x + \alpha)\}$.

Since $v_r(r^*([(\tau_{j+1}, x)]))$ is concave by assumption and $v_a(\tau_j, x + \alpha)$ is concave in x, $S_i(\alpha)$ is convex, i.e. an interval.[14]

Define $S_\infty(\alpha)$ similarly in terms of B_∞ and we get $S_\infty(\alpha)$ convex too.

($S_\infty(\alpha) = \{x : v_a(\tau_j, \Sigma) + \delta v_r(r^*([(\tau_j, \Sigma)])) \leq \delta v_r(r^*([(\tau_{j+1}, x)])) + v_a(\tau_j, x + \alpha)\}$.)

So we have that for $i \in \{0, 1, ...\} \cup \{\infty\}$, $F_i(x) - x$ is quasi-concave. We can see also that since $v_a(\tau_j, x+\alpha)$ is strictly quasi-concave, $F_i(x) - x$ must be strictly quasi-concave.

Let χ_i be the value of x that maximizes $F_i(x) - x$. It follows from quasi-concavity of $F_i(x) - x$ that for $x \leq y \leq \chi_i$, $F_i(y) - F_i(x) \geq y - x$.

Now consider $S_\infty(0)$. $S_\infty(0)$ contains Σ.

Let $l_\infty = inf(S_\infty(0)) \in [-\infty, \Sigma]$ and $h_\infty = sup(S_\infty(0)) \in [\Sigma, \infty)$.

Suppose that $l_\infty < h_\infty$. I will show that in this case y_{i+1} is bounded away from l_∞ eventually.

Define l_i, h_i similarly when $S_i(0) \neq \{\}$.

Take n large enough so that for $m \geq n$:

1a. $S_m(0) \neq \{\}$

1b. $\|F_m - F_\infty\| < \epsilon$. (Uniform metric here as above.)

1c. $|\chi_m - \chi_\infty| < \epsilon$

Where ϵ is chosen such that:

2a. $2\epsilon < \chi_\infty - l_\infty$.

2b. $F_\infty(\chi_\infty - \epsilon) - \chi_\infty > 0$.

[14] These assumptions have been weakened in the main text and there is one point at which this proof needs minor adjustments, to be added shortly.

By condition 2b, if $x \geq \chi_\infty - \epsilon$, $F_\infty(x) > \chi_\infty$ and so by condition 1b, $F_i(x) > \chi_\infty - \epsilon$. So $\max(F_m(x), a^*(\tau_{j+1})) > \chi_\infty - \epsilon$. This implies that if sequence y_i ever leaves the set $(-\infty, \chi_\infty - \epsilon)$, it never returns.

Suppose the sequence y_i remains inside $(-\infty, \chi_\infty - \epsilon)$ for ever. Otherwise y_i has eventual lower bound $\chi_\infty - \epsilon$.

$y_i > x_i$ is a general property of the iterated Riley solution.

We had earlier $F_m(y_m) - F_m(x_m) \geq y_m - x_m$ for $x_m \leq y_m \leq \chi_m$.

$y_{m+1} - x_{m+1} = \max(F_m(y_m), a^*(\tau_{j+1})) - x_{m+1} \geq F_m(y_m) - F_m(x_m) \geq y_m - x_m$ for $x_m \leq y_m \leq \chi_m$.

The first inequality holds because $\max(F_m(y_m), a^*(\tau_{j+1})) \geq F_m(y_m)$ and $F_m(x_m) \geq x_{m+1}$.

$\chi_\infty - \epsilon < \chi_m$ for $m \geq n$.

So for $m \geq n$, $y_m - x_m \geq \delta := y_n - x_n > 0$

Since $x_m \to \Sigma$, y_m is eventually bounded below by $\Sigma + \delta/2 > l_\infty$.

Conclusion 1 *If $l_\infty < h_\infty, y_m$ is bounded below eventually by a lower bound $\underline{\mathbf{b}}$ strictly above l_∞.*

Now we can show that $y_m \to M = \max(h_\infty, a^*(\tau_{j+1}))$

Case 1 *Suppose $l_\infty < h_\infty$.*

Let $G_i(x) := \max(F_i(x), a^*(\tau_{j+1}))$ for $i \in \{0, 1, ...\} \cup \{\infty\}$.

F_∞ on the set $[\underline{b}, a_{\max}]$ has $F_\infty > x$ below M and $F_\infty < x$ above M and no other fixed points.

M is a global attractor and so the sequence defined by $y_{m+1} = G_m(y_m)$ converges to M since $G_m \to G_\infty$ and y_m remains in $(\underline{b}, a_{\max})$.

Case 2 *Now suppose $l_\infty = h_\infty$*

Then $l_\infty = h_\infty = \Sigma$ since $\Sigma \in S_\infty(0)$.

Let $M = \max(h_\infty, a^*(\tau_{j+1})) = \max(\Sigma, a^*(\tau_{j+1}))$ as before.

On $[\Sigma, a_{\max}]$ G has only one fixed point M, and above this $G(x) < x$.

Take $\epsilon > 0$. Let $\inf_{x \in [M+\epsilon, a_{\max}]}(x - G(x)) = \delta$.

Choose n such that for $m \geq n$, $\|G_m - G_\infty\| < \delta$.

When $y_m \leq M + \epsilon$, $y_{m+1} = G_m(y_m) \leq G_m(M + \epsilon) < G(M + \epsilon) + \delta \leq M + \epsilon$

So if y_m reaches the set $(-\infty, M + \epsilon]$, it stays there.

And above this set, y_m decreases by at least δ each time. So y_m remains in the set $(-\infty, M + \epsilon]$ eventually.

We also know that $y_m > x_m \to \Sigma$, and $y_m \geq a^*(\tau_{j+1})$, so $y_m > \Sigma - \epsilon$ eventually .

And $y_m > a^*(\tau_{j+1})$ always. So $y_m > M - \epsilon$ eventually.

Since ϵ is arbitrary, y_m converges to M.

QED ■

Proof. Proof of Proposition 4

ϵ and η as in the statement.

On $\overline{T}$, $a_S(\delta_1)(\tau) > a^*(\tau)$. Since both a_S and a^* are continuous and $\overline{T}$ compact, we can take ϵ_1 such that $a_S(\delta_1)(\tau) > a^*(\tau')$ for $|\tau' - \tau| < \epsilon_1$.

Take $\epsilon \leq \epsilon_1$.

Claim 3 *$\Sigma(\tau) > a^*(\tau)$ for $\tau \in T = \{\tau_0, ...\tau_n\}, \tau > \tau_0$.*

By corollary (N), for $j > 0$, $\Sigma_1(\tau_j) \geq \overline{a}_S(\delta_1)(\tau_{j-1}, \Sigma_1(\tau_{j-1})) \geq a_S(\delta_1)(\tau_{j-1})$.

So $\Sigma_1(\tau_j) \geq a_S(\delta_1)(\tau_{j-1}) > a^*(\tau_j)$ since $|\tau_j - \tau_{j-1}| < \epsilon \leq \epsilon_1$. This proves the claim.

So we know $\Sigma_1(\tau_0) = a^*(\tau_0)$, and for each j, $\Sigma_1(\tau_{j+1}) = \overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j))$.

$\overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j)) = \Sigma_1(\tau_j)$ when $\Sigma_1(\tau_j) \geq a_S(\delta_1)(\tau_j)$. So on $\{\tau_j : a_S(\delta_1)(\tau_j) \leq \overline{a}_S(\delta_1)(\tau_0, a^*(\tau_0))\}$, $\Sigma_1(\tau_j) = \overline{a}_S(\delta_1)(\tau_0, a^*(\tau_0))$. Let the last such τ_j be τ_α.

Take $\eta_2 < \eta/2$ such that for $a_S(\delta_1)(\tau) - \eta_2 \leq x \leq a_S(\delta_1)(\tau)$, $\overline{a}_S(\delta_1)(\tau, x) < a_S(\delta_1)(\tau) + \eta/2$. This is possible because as $\overline{a}_S(\delta_1)(\tau, a_S(\delta_1)(\tau) - \eta_2)$ as a function of τ is continuous, is increasing in η_2 and converges pointwise to $a_S(\delta_1)(\tau)$ as $\eta_2 \to 0$, so converges uniformly to $a_S(\delta_1)(\tau)$ as $\eta_2 \to 0$.

Now take ϵ_2 such that for $|\tau' - \tau| < \epsilon_2$, $|a_S(\delta_1)(\tau') - a_S(\delta_1)(\tau)| < \eta_2$. This is possible because $a_S(\delta_1)$ is continuous and so uniformly continuous on $\overline{T}$.

Suppose that $a_S(\delta_1)(\tau_j) - \eta_2 \leq \Sigma_1(\tau_j) < a_S(\delta_1)(\tau_j) + \eta/2$. Assume that $\epsilon < \epsilon_1, \epsilon_2,$.

If $\Sigma_1(\tau_j) \geq a_S(\delta_1)(\tau_j)$, $\Sigma_1(\tau_{j+1}) = \overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j)) = \Sigma_1(\tau_j)$, so $a_S(\delta_1)(\tau_{j+1}) - \eta_2 \leq a_S(\delta_1)(\tau_j) \leq \Sigma_1(\tau_j) = \Sigma_1(\tau_{j+1}) < a_S(\delta_1)(\tau_j) + \eta/2 < a_S(\delta_1)(\tau_{j+1}) + \eta/2$.

If $\Sigma_1(\tau_j) < a_S(\delta_1)(\tau_j)$ and $a_S(\delta_1)(\tau_{j+1}) - \delta_2 < a_S(\delta_1)(\tau_j) < \Sigma_1(\tau_{j+1}) = \overline{a}_S(\delta_1)(\tau_j, \Sigma_1(\tau_j)) < a_S(\delta_1)(\tau_j) + \delta/2 < a_S(\delta_1)(\tau_{j+1}) + \eta/2$.

So $a_S(\delta_1)(\tau_{j+1}) - \eta_2 \leq \Sigma_1(\tau_{j+1}) < a_S(\delta_1)(\tau_{j+1}) + \eta/2$.

Let $S'(\tau) = \max(\overline{a}_S(\delta_1)(\tau_0, a^*(\tau_0)), a_S(\delta_1)(\tau))$.

It follows that $|\Sigma_1(\tau_j) - S'(\tau_j)| < \eta/2$ for $j > 0$.

Take $\epsilon_3 \leq \epsilon_1, \epsilon_2$ small enough so that $|\overline{a}_S(\delta_1)(\tau, a^*(\tau)) - \overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))| < \eta/2$ for $|\tau - \tau_{\min}| < \epsilon_3$. Then $|S' - S| < \delta/2$.

So for $\epsilon < \epsilon_3$, $|\Sigma_1(\tau_j) - S(\tau_j)| < \eta$. QED ■

Proof. Proof of Proposition 5 (sketch)

1. The differential equation is solvable and has a solution that lies strictly above the discounted Stackelberg curve. Let the difference be at least ϵ.

Note that above ϵ above the discounted Stackelberg curve $\frac{\delta}{\delta a} v_S(\delta_1)$ is bounded above away from 0: $\frac{\delta}{\delta a} v_S(\delta_1)(\tau, a) \leq k < 0$ for $a \geq a_S(\delta_1)(\tau) + \epsilon$.

2. Consider the process starting at $\Sigma_1(\tau_0) = \overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))$ and generated by:

$v_a(\tau_j, \Sigma_1(\tau_{j+1})) + \delta_1 v_r(r^*([(\tau_{j+1}, \Sigma_1(\tau_{j+1}))])) = v_S(\delta_1)\,(\tau_j, \Sigma_1(\tau_j))$.

This is equivalent to:

$[v_S(\tau_j, \Sigma_1(\tau_{j+1})) - v_S(\delta_1)\,(\tau_j, \Sigma_1(\tau_j))] + \delta_1[v_r(r^*([(\tau_{j+1}, \Sigma_1(\tau_{j+1}))])) - v_r(r^*([(\tau_j, \Sigma_1(\tau_{j+1}))]))] = 0$

3. Using the intermediate value theorem, write this as: $(\Sigma_1(\tau_{j+1}) - \Sigma_1(\tau_j))\frac{\delta}{\delta a} v_S(\tau_j, \alpha) + \delta_1(\tau_{j+1} - \tau_j)\frac{\delta}{\delta \tau}[v_r(r^*([(\beta, \Sigma_1(\tau_{j+1}))]))]$, where $\alpha \in [\Sigma_1(\tau_j), \Sigma_1(\tau_{j+1})]$ and $\beta \in [\tau_j, \tau_{j+1}]$.

It follows that Σ_1 above τ_0 has uniform Lipschitz constant $K = \frac{\max(\frac{\delta}{\delta \tau}(v_r(r^*([.]))}{k}$, assuming it remains ϵ above the discounted Stackelberg curve.

4. Rewrite the equation above as:

$\frac{\Sigma_1(\tau_{j+1}) - \Sigma_1(\tau_j)}{\tau_{j+1} - \tau_j} = \delta_1 \frac{\frac{\delta}{\delta \tau}[v_r(r^*([(\beta, \Sigma_1(\tau_{j+1}))]))]}{\frac{\delta}{\delta a} v_S(\tau_j, \alpha)} \in Conv[\delta_1 \frac{\frac{\delta}{\delta \tau}[v_r(r^*([(\tau, a)]))]}{\frac{\delta}{\delta a} v_S(\tau, \alpha)}]$

5. Now consider the solution to the differential equation G:

$\frac{G_1(\tau_{j+1}) - G_1(\tau_j)}{\tau_{j+1} - \tau_j} \in Conv[\delta_1 \frac{\frac{\delta}{\delta \tau}[v_r(r^*([(\tau, a)]))]}{\frac{\delta}{\delta a} v_S(\tau, \alpha)}]$, where this is taken over $(\tau, a) \in [\tau_j, \tau_{j+1}] \times [\Sigma_1(\tau_j), \Sigma_1(\tau_{j+1})]$.

6. The range of this convex hull tends to zero uniformly as the distance between types tends to zero. Take any $\zeta > 0$, then choose $\eta_2 > 0$ such that when all types are within η_2 of each other, the convex hull above has range at most $\zeta < \frac{\epsilon}{2(\tau_{\max} - \tau_{\min})}$.

Therefore $(\Sigma_1(\tau_j) - \Sigma_1(\tau_0)) - (G(\tau_{j+1}) - G(\tau_0)) \leq 2\zeta(\tau_j - \tau_0)$, as long as Σ_1 remains above ϵ above the discounted Stackelberg curve.

7.Combining 3 and 6, the solution remains above ϵ above the discounted Stackelberg curve.

8. $\Sigma_1(\tau_1)$ is arbitrarily close to $\overline{a}_S(\delta_1)(\tau_{\min}, a^*(\tau_{\min}))$ for η_2 small enough, which is arbitrarily close to $G_1(\tau_1)$.

The result follows. ■

References

[1] Dilip Abreu & Faruk Gul, "Bargaining and Reputation," Econometrica, vol. 68(1), pages 85-118, January 2000.

[2] Abreu, Pearce, "Reputational Wars of Attrition with Complex Bargaining Postures", 2006 mimeo

[3] Amaya, An Evolutionary Analysis of Equilibrium Selection in Signaling Games, working paper 2007

[4] In-Koo Cho. A refinement of sequential equilibrium. Econometrica, 55(6):1367-89, 1987.

[5] I. K. Cho and D. Kreps. Signaling games and stable equilibria. Quarterly Journal of Economics, 56:601-612, 1987

[6] Cho, In-Koo & Sobel, Joel, "Strategic stability and uniqueness in signaling games," Journal of Economic Theory, vol. 50(2), pages 381-413, April 1990

[7] Martin Cripps, Eddie Dekel and Wolfgang Pesendorfer, "Reputation with Equal Discounting in Games with Strictly Conflicting Interests", Journal of Economic Theory 2005.

[8] M. Celentani, D. Fudenberg, D.K. Levine, and W. Pesendorfer. Maintaining a reputation against a long-lived opponent. Econometrica 1966.

[9] Cripps, Martin W. & Mailath, George J. & Samuelson, Larry, 2007. "Disappearing private reputations in long-run relationships," Journal of Economic Theory, Elsevier, vol. 127(1), pages 287-316.

[10] Evans, R. and J.P. Thomas, (1997), Reputation and Experimentation in Two-Player Repeated Games with Discounting, Econometrica, 65, 1153-1173

[11] Drew Fudenberg and David Levine. Reputation and equilibrium selection in games with a patient player. Econometrica, 57:759-778, 1989.

[12] Drew Fudenberg and David Levine, Maintaining a Reputation when Strategies are Imperfectly Observed, RES 1992

[13] Drew Fudenberg & Jean Tirole, "A "Signal-Jamming" Theory of Predation," RAND Journal of Economics, The RAND Corporation, Volume 17, No. 3. Issue: Autumn 1986

[14] Holmstrom, B. "Managerial incentive problems: A dynamic perspective." Review of Economic Studies 66: 169-182, Jan 1999 17(3), pages 366-376.

[15] Ayca Kaya, "Repeated Signaling Games", working paper (August 2007). Available at SSRN: http://ssrn.com/abstract=765905

[16] Kohlberg, Elon & Mertens, Jean-Francois, "On the Strategic Stability of Equilibria," Econometrica, vol. 54(5), pages 1003-37, September 1986

[17] David Kreps and Robert Wilson, "Reputation and imperfect information", Journal of Economic Theory, 27:253-279, 1982.

[18] George J. Mailath, Masahiro Okuno-Fujiwara, and Andrew Postlewaite. Belief based refinements. Journal of Economic Theory, 60:24-276, 1993.

[19] Mailath,G.J. & Samuelson,L., 1998."Your reputation is who you're not, not who you'd like to be", Working papers 18, Wisconsin Madison - Social Systems

[20] Mailath, George J & Samuelson, Larry, "Who Wants a Good Reputation?", Review of Economic Studies, 415-41, April 2001

[21] Noldeke, G. & Samuelson, L., 1995. "A Dynamic Model of Equilibrium Selection in Signaling Markets," Working papers 9518, Wisconsin Madison - Social Systems

[22] Ramey, D1 Signalling Equilibria with Multiple Signals and a Continuum of Types, Journal of Economic Theory, 508-531, May 1996

[23] Joel Sobel. Signaling games. Technical report, UCSD, 2007

[24] Michael Spence, "Job Market Signaling". The Quarterly Journal of Economics, Vol. 87, No. 3., Aug., 1973