ISSN 1471-0498

DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

Fooling Some of the People Some of the Time: Reputation Management and Optimal Betrayal

Andrew Mell

Number 770
December 2015

Manor Road Building, Oxford OX1 3UQ

Fooling Some of the People Some of the Time: Reputation Management and Optimal Betrayal

Andrew Mell*
andrew.mell@economics.ox.ac.uk

October 29, 2015

"On the whole human beings want to be good, but not too good, and not quite all the time."

George Orwell, *All Art is Propaganda: Critical Essays.*

Abstract

A rational long lived player plays against a series of short lived players who use a variant of the Adaptive Play behavioral rule. In equilibrium, under certain conditions, there will be a cut-off level of reputation. If their reputation is below the cut-off, they will build their reputation, and consume out of their reputation if it is above the cut-off. Over the long run, their reputation oscillates around the cut-off. A public relations professional can manipulate the sampling of the short lived players to the benefit of the long lived player. As a result a patient long lived player's behavior will worsen while an impatient long lived player's behavior will improve.

Keywords: Reputation, Adaptive Play, Monitoring, Expectation Formation
JEL Classification: D82, D83, D84

*Corpus Christi College, Merton Street, Oxford, OX1 4JF, UK. This paper has been based on work conducted for a chapter of my PhD Thesis, and so thanks are due to my supervisors, Peyton Young and Meg Meyer for their helpful insights. Thanks are also due to Colin McDiarmid who provided some helpful pointers for some of the mathematical issues. Particular thanks are also due to Bassel Tarbush whose advice around dynamic optimization was invaluable. Naturally all remaining errors are my own.

1 Introduction and Relevant Literature

Current models of reputation, imbue short lived players with remarkable capacities for gathering and analyzing data. These models also frequently rely on players pursuing extremely complicated mixed strategies. There should be room to consider that short lived players in a reputation game might behave in a boundedly rational manner. After all, they play the game once and then move on to other things in their lives, deploying their fullest deductive faculties to this one interaction would seem somewhat irrational.

Furthermore, the public relations industry is dedicated to helping their clients improve their reputation.[1] Yet they do not control their clients' actions, which is the only way to improve a reputation in standard models. They attempt to control which actions people have heard about. Reputation models in which all short lived players already know all the past moves of players building a reputation leave no room for such manipulation of the information set.

This paper draws on boundedly rational behavior models used in evolutionary game theory to construct a new way of thinking about the behavior of short lived players in reputation models. However, the long-lived player continues to be treated as a rational *homo economicus*. Specifically we adapt the model of 'adaptive play' from Young (1993, 1998) and consider an environment where short lived players best respond to a sample of some actions taken by the long lived player.[2] This paper considers two possibilities where the sampling process is 'unmanaged', and where it is 'managed' by a public relations (PR) specialist who skews the sampling in favor of the long lived player.

When reputation is unmanaged, the optimal strategy of the long lived player depends crucially on how patient they are, and whether their temptation to betray is increasing or decreasing in the probability with which they expect to be trusted. Whether the temptation to betray is increasing or decreasing, there is the possibility of extremely impatient long lived players who never engage in reputation building and extremely patient long lived players who do little other than build reputations.

The interesting cases concern long lived players with an intermediate level of patience. Where the temptation to betray is increasing in the probability

[1] See for example, the Chartered Institute of Public Relations website. Answering the question, 'what is PR?' it says 'Effective PR can help manage reputation…' `http://www.cipr.co.uk/content/careers-cpd/careers-advice-and-case-studies/what-pr`.

[2] Note that Kandori, Mailath and Rob (1993) have a model similar to that in Young (1993, 1998).

of being trusted, the optimal strategy involves building up a reputation when their reputation is poor, but consuming out of good reputations. Over the long run, such players gravitate towards an intermediate reputation which will be sufficient to persuade some, but not all short lived players they encounter to trust them. Where the temptation to betray is decreasing in the probability of being trusted, long lived players have a cut-off. They will build up reputations that are already above a certain threshold, but they will consume out of reputations that are below a certain threshold. So depending on how their reputation starts, they will either end up with a perfect reputation, jealously guarded, or no reputation at all in the long run.

The management of reputation broadly speaking has two effects. First, even a low stock of good behavior is sufficient to ensure the long lived player will be trusted with a very high probability. This essentially lowers the cost of building a good reputation, so less patient players become willing to build reputations, leading to more good behavior on the part of the long lived player. However, a smaller number of instances of good behavior are now required in order to ensure that they will be trusted, so those players who do build reputations, will behave less well and be trusted more frequently over the course of the game.

In examining a reputation game where the short lived players do not behave as the standard rational maximizers of economics, this paper departs from the incomplete information approach to reputations developed in Kreps and Wilson (1982); Milgrom and Roberts (1982); and Kreps et al. (1982), and generalized in Fudenberg and Levine (1989, 1992). In doing so, it bears similarities to Selten's attempt to resolve the Chain Store Paradox in Selten (1978), which proposed different nested levels of decision making, falling short of the rational calculations of *homo economicus*.

With rational short lived players, reputation models can be constructed along the lines of the incomplete information approach with fully rational players and these models do achieve some intuitive results. However we should remember that the equilibrium strategies in these models can be extremely complicated and rely on short lived players being able to gather and analyze vast amounts of information. Even if they had the capacity to do this, it is not certain that it would be rational to do so for an interaction in which they will take part once before moving on to other things in their lives. Hence the interest in a model where short lived players behave with bounded rationality. The similarity in the situation of short lived players in a reputation game and adaptive players makes adaptive play a natural model to start with.

In building a model that allows the normal *homo economicus* of traditional models to interact with their boundedly rational cousins from evolutionary game theory, this model is similar to that in Ellison (1997). That paper examines the circumstances under which the 'one rational guy' in a field of players using the fictitious play rule might seek to shift the game from one equilibrium to another. Jehiel and Samuelson (2012), using the Analogical Based Expectation concept from Jehiel (2005), build a reputation model in which the short lived players conduct traditional Bayesian inference. However, their inferences are based on a mistaken model which sees all types of long lived players using a stationary (possibly mixed) strategy. The authors comment that the results of their model are similar to those that would result from one in which a rational long lived player played against behavioral players employing the fictitious play rule. That model can have no room for public relations in managing reputations, and cannot produce reputations that will be strong enough to convince some short lived players, but not others. Both results come naturally to the model developed here. Although interactions between rational players and adaptive players have been considered before in Sáez-Martı and Weibull (1999), that concerned a bargaining context rather than a reputation building one.

Previous papers have considered the possibility of limiting what the short lived players know about the past behavior of the long lived player in reputation games. For example, Liu (2011); and Liu and Skrzypacz (2014) consider reputation models where short lived players have access only to limited records. These papers still assume short lived players have impressive information processing capabilities, and the nature of their limited information still leaves little room for public relations. Furthermore the results presented here provide an interesting contrast to previous work on reputation with limited records. In those models, reputation cycles or "bubbles" develop naturally, but the length of these cycles increases with the size of the bounded memory. In contrast, in the model developed here, for a wide range of players and parameters, increasing the amount of information will actually lead to less reputation building by the long lived player.

In a similar vein, Mailath and Samuelson (2001); Phelan (2006); and Wiseman (2008) have considered the possibility that not all of the information gathered by short lived players is relevant. There might be some probability of a hidden change in the type of the long lived player between periods. Both the bounded memory models discussed above, and these models with hidden type changes are presented as a resolution to a potential problem highlighted in Cripps, Mailath and Samuelson (2004) which shows that reputation can only be a short run phenomenon. In the long run, the type of the long lived player must be revealed. Hidden type changes and bounded memory ensure this does not happen either because uncertainty is periodically refreshed, or because short lived players never have enough information to be too certain. The type revelation issues highlighted in Cripps, Mailath and Samuelson (2004) are not an issue in this model because of the behavioral rules of the short lived players.

While this paper does not deal with advertising *per se*, the field of advertising is related to the field of public relations. There is a wide literature on advertising in economics, primarily concerned with distinctions between informative or persuasive advertising.[3] Nelson (1974) suggests that better brands will advertise more. This is generally because, in his model, the better a brand, the more they will benefit from the extra sales generated by advertising. In this paper, the second effect of PR specialists mentioned above points in the other direction. When a long-lived player makes use of public relations services, they can enjoy a better reputation without actually having to take the action for which they desire a reputation so frequently. As a result, the frequency with which they cheat on their reputation increases.

The next section describes and outlines the model. Section 3 outlines the optimal strategy for the long lived player when there is no effort to manage the long-lived player's reputation. Section 4 compares the results in this model with the results in other models of reputation. Section 5 considers what the effect might be of reputation management by a public relations expert and how that might affect the behavior of the long lived player. Finally, Section 6 concludes and highlights several avenues for future research.

2 The Model

We consider a situation in which the same player 1 plays the same stage game indefinitely against a series of players each of whom plays in the position of player 2 for one period only. We assume that player 1's time preferences are captured by the discount factor, δ. Since player 2 only plays in one period, their time preferences are not relevant.

In the stage game player 1 and player 2 simultaneously choose between two actions. For player 1, these actions are T and B. For player 2, they are L and R. The payoffs are as shown in Table 1. For player 1 we assume that $a > b > c > d$, and for player 2 we assume that $x > z$ and $y > w$, we also assume $x > y$, so that the outcome (T,L) Pareto dominates the outcome (B,R).

For player 1, the action T represents the exertion of costly effort, increasing

[3] An excellent survey of this literature can be found in Bagwell (2007).

		Player 2	
		L	R
Player 1	T	b, $\underline{x}$	d, z
	B	$\underline{a}$, w	$\underline{c}$, $\underline{y}$

Table 1: The payoffs in the trust game, where $a > b > c > d$, $y > w$, $x > z$, and $x > y$.

the surplus generated by the relationship. As such, the action B dominates the action T. For player 2, the action L represents trusting player 1, while the action R represents a refusal to trust player 1. L is the best response to T and R is the best response to B. This is sometimes interpreted as a product choice game. T represents the choice to produce a high quality product, while B represents the choice to produce a low quality product. L represents a large order while R represents a small order. This game is frequently used in the reputation literature because the only Nash equilibrium is (B,R), but it is Pareto dominated by (T,L). From the perspective of player 1, the question then becomes how can they persuade player 2 that they are going to play T, and how frequently do they have to pay the opportunity cost of actually playing T in order to do so?

First, consider the game from the perspective of player 2. Suppose that player 2 believes that player 1 will play T with probability p. Then they will trust player 1 and play L if and only if:

$$px + (1-p)\,w > pz + (1-p)\,y \Leftrightarrow p > \frac{y-w}{(y-w)+(x-z)}. \qquad (1)$$

The question of expectation formation then becomes crucial. In this paper, we will consider the consequences of a simple rule for expectation formation. Suppose that there exists a 'social memory' of m actions that the long lived player has taken in the past. We assume that $m \geq 3$. When player 2 is called to move, they sample one of the actions from the social memory selected at random, and believe that this is the action player 1 will take when playing the stage game against them. This means they will either form the belief $p = 1$, if they happen to sample a T, or $p = 0$, if they happen to sample a B. So if they sample a T, they will play L and if they sample a B, they will play R. Essentially, player 2 best responds to the precedent they have heard of, and the social memory represents the set of precedents they might possibly have heard of.

The evolution of the social memory is important. It is assumed that the action played by player 1 in period t is added to the social memory in period $t+1$. In doing so, it replaces one of the m actions that were in the social memory in period t. The action it replaces is selected at random, with all actions in the social memory in period t having an equal chance of being replaced.

At this point it is worth commenting on the similarities and differences between the short lived players in this model and adaptive players in the evolutionary model of Young (1993). Adaptive players, when they are called to move know the actions taken by the person playing in the position of their opponent in s of the last m interactions. Those s interactions have been sampled at random without replacement.

So the first difference between adaptive players in evolutionary game theory and short-lived players in this model is that, in this model, players are limited to a sample of size $s = 1$. While expanding the model to the case of $s > 1$ remains an area of ongoing research, the case of $s = 1$ has a neat interpretation and is deserving of particular attention. Every reader will doubtless be able to think of an occasion where they have made a decision such as where to go on holiday, which smartphone or laptop to buy, largely based on a single recommendation. It may be that only one recommendation was received, or it might be that one recommendation, because of the way it was phrased, received far more weight than any others. It would seem to be a fallibility of human beings that we are far too prone to being persuaded by anecdote rather than data.[4,5]

The second difference concerns the evolution of the social memory from which instances are sampled. In normal adaptive play models, this is simply the m previous actions. So when a new action is taken, it is the oldest action that leaves the social memory. In this model, the action that leaves the social

[4]The author's wife still complains to the author about an incident before they were dating where she chose which mobile phone network to join based on his recommendation, only to discover that she had no signal in her own halls of residence. On a less personal note, British politicians during the 2015 Election Campaign frequently justified their policies by talking about their encounters with an ordinary voter they had met.

[5]A highly topical example at the time of writing concerns UK public attitudes to refugees from Syria dying as they attempt to reach Europe by crossing the Mediterranean in unsafe boats. While numbers and statistics on the issue had been prevalent in the press for six months, the issue had barely registered on the public consciousness. However on 2nd September 2015, a photograph of three year old Aylan Kurdi's drowned body having washed up on the beach in Turkey went 'viral' on Social Media. The issue then exploded, and a petition asking the UK government to accept more refugees gained more than 300,000 signatures in less than 48 hours.

memory to make way for a new action is selected at random. This matches intuitions about reputation that there is a degree of randomness about which actions are remembered and available to be sampled through a process of asking around, and which instances people forget about quickly.[6]

So, for player 1, the optimal strategy now takes on the characteristics of an optimal control problem. Let t, be the number of instances of T in the social memory. We will refer to this as the 'state'. The state space allows t to take any integer value in the interval $[0, m]$.

From any state, the play of B leads to a better instantaneous payoff, but risks moving down to a lower state. The play of T leads to a lower instantaneous payoff, but has a chance of moving to a higher state. The benefit of higher states is that the short lived player is more likely to sample a T rather than a B and so more likely to trust the long lived player 1 in a higher state. In this sense, player 1's reputation is stronger, the higher the value of t.

We can write the Bellman equation for this optimal control problem as follows:

$$V(t) = \max \left\{ \begin{array}{c} \frac{t}{m}b + \left(1 - \frac{t}{m}\right) d + \delta \frac{t}{m} V(t) + \delta \left(1 - \frac{t}{m}\right) V(t+1), \\ \frac{t}{m}a + \left(1 - \frac{t}{m}\right) c + \delta \frac{t}{m} V(t-1) + \delta \left(1 - \frac{t}{m}\right) V(t) \end{array} \right\} \tag{2}$$

The top line represents the expected payoff from playing T. The first two terms give the expected instantaneous utility payoff from playing T in state t, while the second two terms give the expected continuation value from the next period onwards as a result of any change in state coming about from playing T. The second line has a similar interpretation, but for the play of B. In state t, player 1 has a choice between playing T or B, and so the continuation value to being in the state is whichever is the larger of these two expressions.

3 Unmanaged Reputation

In this section, we find the equilibrium in the game by finding the solution to player 1's optimal control problem. Notationally, let $\mathbf{D}V(t) = V(t) -$

[6] It also has the mathematically convenient feature of collapsing the state space. Suppose there are two states, 1 and 2, with the same number of instances of T and B. Suppose that in state 1, the oldest action was T and in state 2 the oldest action was B. If actions exit in sequence, then these two states are not strategically equivalent. Taking the action T in state 1 will move to a state with the same number of Ts and Bs. Taking the action T in state 2 will move to a state with a larger number of Ts. However if actions exit the history at random, then the transition probabilities depend only on the number of Ts and Bs available to drop out of the state and these two states would be strategically equivalent.

$V(t-1)$ represent the difference in the value function between states t and $t-1$. Intuitively, $\mathbf{D}V(t) > 0$, as a higher chance of being trusted leads to higher payoffs. We can also say, from (2) that the optimal action in any state t will be T if and only if:

$$\delta \frac{t}{m} \mathbf{D}V(t) + \delta \left(1 - \frac{t}{m}\right) \mathbf{D}V(t+1) > \frac{t}{m}(a-b) + \left(1 - \frac{t}{m}\right)(c-d). \quad (3)$$

The left hand side represents the difference in future asset values of player 1's reputation as a result of playing T rather than B. The right hand side represents the difference in instantaneous payoff as a result of playing B rather than T. Playing T is better if the gain in asset value of player 1's reputation from playing T outweighs the gain in instantaneous payoff from playing B.

It is convenient to define the following functions of t:

$$H(t) = \delta \frac{t}{m} \mathbf{D}V(t) + \delta \left(1 - \frac{t}{m}\right) \mathbf{D}V(t+1), \quad (4)$$

$$G(t) = \frac{t}{m}(a-b) + \left(1 - \frac{t}{m}\right)(c-d). \quad (5)$$

Then (3) can be expressed as

$$H(t) > G(t). \quad (6)$$

The solution turns out to depend crucially on whether player 1's temptation to cheat on their reputation is increasing or decreasing in the strength of their reputation. The former case is characterized by parameters such that $a-b > c-d$. The latter case is characterized by parameters such that $a-b < c-d$.

3.1 Increasing Temptation to Cheat

As well as depending on whether the temptation to cheat on a reputation becomes stronger as one's reputation improves, the optimal strategy for player 1 also depends on how patient they are. There are three possibilities. First, they are so impatient, that they never build any kind of reputation. Second, they are so patient, they they always build up their reputation to the point of perfection. Finally at intermediate levels of patience, there will be some reputation building, but also some consumption out of their reputation. Each case is dealt with in turn.

3.1.1 No Reputation

Suppose player 1 is insufficiently patient to ever build a reputation. Then the value function at any state can be written as:

$$V(t) = c + \frac{t}{m}(a-c) + \delta \frac{t}{m} V(t-1) + \delta \left(1 - \frac{t}{m}\right) V(t). \tag{7}$$

So the change in the value function, $\mathbf{D}V(t)$ can be written as:

$$\mathbf{D}V(t) = \frac{\frac{1}{m}(a-c) + \delta \frac{t-1}{m} \mathbf{D}V(t-1)}{1 - \delta\left(1 - \frac{t}{m}\right)}. \tag{8}$$

Lemma 1 demonstrates that this must be a constant:

Lemma 1 *When player 1 always plays B between state* 0 *and state* t^**, the change in the value function,* $\mathbf{D}V(t)$ *does not depend on the state,* t $\forall t \in \{1, 2, \ldots t^*\}$*, and for all these values is equal to:*

$$\mathbf{D}V(t) = \frac{\frac{1}{m}(a-c)}{1 - \delta + \delta/m} \tag{9}$$

The proof works by induction and has been confined to Appendix A on account of its length. It is however worth noting here that nothing in the proof requires that $a - b > c - d$, so this Lemma will also be relevant in the case where there is a decreasing temptation to cheat.

If the growth in the value function is constant and does not depend on t, then the function $H(t)$ which is a weighted average of the growth in the value function at two points will also be constant and will not depend on t. So, for all values of t,

$$H(t) = \frac{\frac{\delta}{m}(a-c)}{1 - \delta + \delta/m}.$$

It is now straightforward to find the circumstances under which it is always optimal to play B, and these are contained in Proposition 2

Proposition 2 *If the temptation to cheat is stronger for a player more likely to be trusted,* $a - b > c - d$*, then the strategy of always playing B will indeed be optimal provided:*

$$\delta < \delta^* = \frac{c-d}{\left(1 - \frac{1}{m}\right)(c-d) + \frac{1}{m}(a-c)}. \tag{10}$$

Proof. In order to confirm that the hypothesized strategy is indeed optimal, we need to ensure that $\forall t$, $G(t) > H(t)$. Under the candidate strategy profile, $H(t)$ is constant in t and its value is known. Where $a - b > c - d$, the smallest value of $G(t)$ occurs at $t = 0$. So if $G(0) > H(0)$, then the strategy is optimal $\forall t$. This requires:

$$c - d > \frac{\frac{\delta}{m}(a - c)}{1 - \delta + \delta/m},$$

$$\Rightarrow \delta < \delta^* = \frac{c - d}{\left(1 - \frac{1}{m}\right)(c - d) + \frac{1}{m}(a - c)} \tag{10}$$

■

3.1.2 Perfect Reputation

If player 1 plays T all the time, then the value function will take the form:

$$V(t) = d + \frac{t}{m}(b - d) + \delta \frac{t}{m} V(t) + \delta \left(1 - \frac{t}{m}\right) V(t + 1). \tag{11}$$

The change in the value function can then be expressed as:

$$\mathbf{D}V(t) = \frac{\frac{1}{m}(b - d) + \delta \left(1 - \frac{t}{m}\right) \mathbf{D}V(t + 1)}{1 - \delta \frac{t-1}{m}}. \tag{12}$$

Lemma 3, which is similar to Lemma 1 above, shows that (12) must be constant under these circumstances.

Lemma 3 *When player 1 always plays T between state t^* and state m, the change in the value function, $\mathbf{D}V(t)$ does not depend on the state, t $\forall t \in \{t^* + 1, t^* + 2 \ldots m\}$, and for all these values is equal to:*

$$\mathbf{D}V(t) = \frac{\frac{1}{m}(b - d)}{1 - \delta + \delta/m} \tag{13}$$

The proof has been confined to Appendix A, but follows precisely the same steps *mutatis mutandis* as did the proof of Lemma 1. As such, it is once more the case that nothing in the proof requires that $a - b > c - d$. So Lemma 3 will also be relevant in the case where there is a decreasing temptation to cheat.

The invariance of $\mathbf{D}V(t)$ with t, just as before, feeds through to an invariance of $H(t)$ to t. So that $\forall t$,

$$H(t) = \frac{\frac{\delta}{m}(b - d)}{1 - \delta + \delta/m}. \tag{14}$$

Proposition 4 confirms that for a sufficiently large, δ, the strategy of always playing T is optimal.

Proposition 4 *If the temptation to cheat is stronger for a player more likely to be trusted, $a-b>c-d$, then the strategy of always playing T will indeed be optimal provided:*

$$\delta > \delta^{**} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)}. \tag{15}$$

Proof. In order for player 1 to optimally always play T, it must be the case that $H(t) > G(t)$ holds $\forall t$. By Lemma 3, $H(t)$ is constant and does not vary with t. Given the increasing temptation to cheat, the highest value for $G(t)$ is where $t=m$. So always playing T will be optimal for player 1 if and only if:

$$\frac{\frac{\delta}{m}(b-d)}{1-\delta+\delta/m} > a-b,$$

$$\Rightarrow \delta > \delta^{**} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)}. \tag{15}$$

■

Before considering optimal strategies in the range $[\delta^*, \delta^{**}]$, it is worth pausing to demonstrate that it must indeed be the case that $\delta^{**} > \delta^*$. To suppose otherwise would in fact contradict the main parameter assumption of this section.

$$\delta^{**} < \delta^* \Leftrightarrow (a-b)(a-c) < (b-d)(c-d)$$

The contradiction is clear, since we know that $a-b>c-d$ and, by implication $a-c>b-d$.

It can also be shown that $\delta^* < 1$. To suppose that $\delta^* \geq 1$ would imply that $c-d \geq a-c$. But it is known that $a-c > a-b > c-d$, hence supposing $\delta^* \geq 1$ yields a contradiction. So as Player 1 becomes more patient, it becomes impossible for her optimal strategy to involve the continual play of B. As regards δ^{**}, both $\delta^{**} < 1$ and $\delta^{**} > 1$ remain open as possibilities.

3.1.3 Some Reputation Building

We now examine the case where $\delta \in [\delta^*, \delta^{**}]$. In this case, the candidate optimal strategy is to play T $\forall t < t'$, and to play B $\forall t \geq t'$. So there is a reputation building phase at the lower end of the state space and the there is a reputation consuming phase at the upper end of the state space. The state t' is the lowest state in the reputation consuming part of the state space. The nature of this optimal strategy is captured by Proposition 5

Proposition 5 *If $\delta \in [\delta^*, \delta^{**}]$, and $a - b > c - d$, then the optimal strategy for player 1 involves building a reputation and playing T when $t < t'$ and consuming out of their reputation and playing B where $t \geq t'$, where t' is the sole integer in the interval $\left[t'_-, t'_- + 1\right]$, and*

$$t'_- = \frac{\delta (a - c) - m (1 - \delta + \delta/m)(c - d)}{(1 - \delta + 2\delta/m)[(a - b) - (c - d)]} \tag{16}$$

The proof of Proposition 5 is extremely lengthy and requires that several intermediate results be proved *en route*. As such, it has been confined to Appendix B. In essence, it shows that the hypothesised strategy ensures that there is a single crossing point between the functions $H(t)$ and $G(t)$, such that $G(t)$ crosses the $H(t)$ function from below.

As regards t'_-, the formulation:

$$\frac{t'_-}{m} = \frac{\frac{\delta}{m}(a - c) - (1 - \delta + \delta/m)(c - d)}{(1 - \delta + 2\delta/m)[(a - b) - (c - d)]} \tag{17}$$

will frequently be more useful since it shows where t' must lie relative to size of the size of the memory.

So at intermediate levels of patience, player 1 will engage in reputation building. However the reputation they build will not be perfect and will leave some positive probability of "reputational failure" where their reputation is not sufficient to persuade a short lived player to trust them. The pattern of movement through the state space induced by the optimal strategy is shown in Figure 1

Figure 1 shows that player 1 will move up the state space during the reputation building phase and will move down the state space during the reputation consuming phase. Over the long run, player 1 will be oscillating between two states, t', the lowest state of the reputation consuming phase, and $t' - 1$, the highest state of the reputation building phase. What the optimal strategy, and Figure 1 show, is that from wherever player 1 starts in the state space, they will eventually gravitate towards these two states.

Intuitively, as player 1's reputation improves, they are more likely to be trusted by short lived players, so their temptation to cheat gets stronger. This

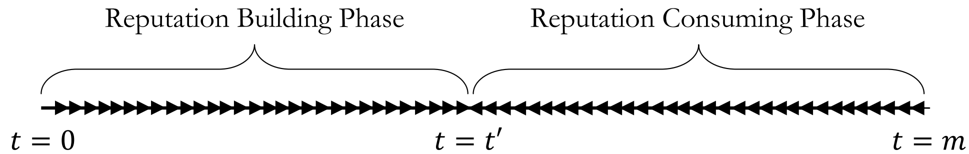

Figure 1: The movement through the state space induced by player 1's optimal strategy.

result has intuitive similarities with the results in Aperjis, Zeckhauser and Miao (2014) where players face a stochastically variable temptation to cheat, except in this case, the temptation to cheat varies systematically as a player moves through the state space. So as their reputation improves, they must be more and more patient in order to forego the increasing benefits of cheating and play T. Since their patience is limited, $\delta \leq \delta^{**}$, eventually, the temptation becomes too strong and they will play B instead.

Further insight into the dynamics of the optimal strategy can be obtained by differentiating these critical values with respect to some of the other parameters reveals how they change as the parameters of the game change. In particular:

$$\frac{d\delta^*}{dm} = \frac{(c-d)\left[(a-c)-(c-d)\right]}{m^2\left[\left(1-\frac{1}{m}\right)(c-d)+\frac{1}{m}(a-c)\right]^2} > 0, \tag{18}$$

$$\frac{d\delta^{**}}{dm} = \left.\frac{(a-b)\left[(b-d)-(a-b)\right]}{m^2\left[\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)\right]^2}\right|_{\delta^{**}<1} > 0, \tag{19}$$

$$\frac{\partial\frac{t'_-}{m}}{\partial\delta} = \frac{a-d}{m\left(1-\delta+2\delta/m\right)^2\left[(a-b)-(c-d)\right]} > 0 \tag{20}$$

$$\frac{\partial\frac{t'_-}{m}}{\partial m} = -\frac{\delta(1-\delta)(a-d)}{m^2\left(1-\delta+2\delta/m\right)^2\left((a-b)=(c-d)\right]} < 0. \tag{21}$$

The result in (20) that $\partial\frac{t'_-}{m}/\partial\delta > 0$ is highly intuitive. As player 1 becomes more patient, she becomes more willing to sacrifice short term gains by cheating on her reputation for long term gains from being trusted. So the reputation building region of the state space expands at the expense of the reputation consuming region of the state space.

The results in (18) and (21) are perhaps best explained by the impact of m on the benefits of transitioning up or down one state in the state space. The larger is m, the smaller is the difference in the probability of being trusted between any two consecutive states. So the marginal incentive to forgo the benefits of opportunistic action by playing T rather than B are smaller. In essence, a coarsening of the possible levels of player 1's reputation (reduction in m) makes it cheaper to build good reputations and so leads to more players willing to build better reputations. (reduction in δ^* and increase in t_-/m).

The sign of (19) is less clear. δ^{**} will increase in m whenever $b - d > a - b$ and will decrease in m whenever $b - d < a - b$. It is also the case that $\delta^{**} < 1 \Leftrightarrow b - d > a - b$. Hence $d\delta^{**}/dm > 0 \Leftrightarrow \delta^{**} < 1$. Since δ^{**} is only of interest when $\delta^{**} < 1$, this is intuitive and similar to the results for δ^*. As m increases, the cost of building a good reputation increases and so the patience required to build a perfect reputation must also increase.

3.2 Decreasing Temptation to Cheat

When the temptation to cheat decreases as a reputation grows stronger, that implies that $c - d > a - b$. The previous analysis assumed the opposite was the case. The structure of the analysis here is very similar to that for the case of an increasing temptation to cheat. This is useful as some of the results from the previous analysis can be reused. Specifically, nothing in the proofs of Lemmas 1 and 3 required that the temptation to cheat be increasing. So these Lemmas will still apply in this case.

There are three cases each of which is similar to one of the cases when there was an increasing temptation to cheat. First, if player 1 is not very patient, it will turn out that they will not build any kind of reputation. If player 1 is extremely patient, they will do nothing but build a reputation. Finally, at intermediate levels of patience, player 1 will build their reputation in some states by playing T, and will consume out of their reputation in other states by playing B. However a key difference is that whereas, previously, when player 1 had an intermediate level of patience, they would play T in the low states, building up their reputation and play B in the high states, consuming their reputation. In this case, they will play B in the low states, finding it too expensive to build any kind of reputation, but they will play T in the higher states, and find it optimal to improve on an intermediate reputation.

Once again, we examine each case in turn, starting with the case where player 1 is insufficiently patient to build any kind of reputation; next moving onto the case where player 1 is so patient, they never do anything but build a reputation. Finally we will examine the intermediate case where some reputation building and some reputation consumption take place.

3.2.1 No Reputation

By Lemma 1, if player 1 always plays B, then $\forall t$,

$$H(t) = \frac{\frac{\delta}{m}(a-c)}{1-\delta+\delta/m}. \tag{22}$$

Proposition 6 follows as a consequence.

Proposition 6 *The strategy of always playing B is optimal provided that:*

$$\delta < \delta' = \frac{a-b}{(a-b)+\frac{1}{m}(b-c)}. \tag{23}$$

Proof. The strategy of always playing B will be optimal if and only if $G(t)$ lies everywhere above $H(t)$. Given $c-d > a-b$, the lowest point on the function $G(t)$ will be where player 1 is completely trusted, $t = m$. By Lemma 1, $H(t)$ is invariant to t. So if $H(m) < G(m)$, then $G(t) > H(t)$, $\forall t$.

$$\frac{\frac{\delta}{m}(a-c)}{1-\delta+\delta/m} < a-b,$$

$$\Rightarrow \delta < \delta' = \frac{a-b}{(a-b)+\frac{1}{m}(b-c)}. \tag{23}$$

■

3.2.2 Perfect Reputation

By Lemma 3, if player 1 plays T $\forall t$, then,

$$H(t) = \frac{\frac{\delta}{m}(b-d)}{1-\delta+\delta/m}. \tag{24}$$

Proposition 7 follows as a consequence.

Proposition 7 *The strategy of always playing T is optimal provided that:*

$$\delta > \delta'' = \frac{c-d}{(c-d)+\frac{1}{m}(b-c)}. \tag{25}$$

Proof. The strategy of always playing T will be optimal if and only if $H(t)$ lies everywhere above $G(t)$. Given $c-d>a-b$, the highest point on the function $G(t)$ will be where player 1 is not trusted, $t=0$. By Lemma 3, $H(t)$ is invariant to t. So if $H(0)<G(0)$, then $H(t)>G(t)$, $\forall t$.

$$\frac{\frac{\delta}{m}(b-d)}{1-\delta+\delta/m} > c-d,$$

$$\frac{\delta}{m}(b-d) > (1-\delta+\delta/m)(c-d),$$

$$\delta > \delta'' = \frac{c-d}{(c-d)+\frac{1}{m}(b-c)}. \tag{25}$$

■

Before moving on to consider the case where there is *some* reputation building, it is worth assuring ourselves that $\delta''>\delta'$. This turns out to indeed be the case, if and only if $c-d>a-b$, which must be the case where the temptation to cheat is decreasing. It can also be shown that $\delta',\delta''<1$, to suppose otherwise would require that $b-c<0$, which clearly runs counter to the restrictions placed on Player 1's payoffs. This presents an interesting contrast between the case of an increasing and a decreasing temptation to cheat. When the temptation to cheat is increasing, there is no guarantee that $\delta^{**}<1$, so there is no guarantee that it is possible to encounter players sufficiently patient to build a perfect reputation. Such a possibility is guaranteed in the case where the temptation to cheat is decreasing.

3.2.3 Some Reputation Building

Now consider the case where $\delta\in[\delta',\delta'']$. In this case, the candidate strategy is to play B in all states, $t<t''$, and to play T in all states $t\geq t''$.

Proposition 8 *Where $\delta\in[\delta',\delta'']$, and $c-d>a-b$, the optimal strategy is to play T $\forall t\geq t''$ and to play B $\forall t<t''$, where t'' is the sole integer in the interval $\left(t''_{-},t''_{-}+1\right)$, where*

$$t''_{-} = \frac{m(1-\delta)(c-d)-\delta(b-c)}{(1-\delta)\left[(c-d)-(a-b)\right]}. \tag{26}$$

The proof has been confined to Appendix D.

The optimal strategy for player 1 in these circumstances involves a critical state of their reputation. Above this state, they find it optimal to build a reputation and continue to build that reputation until it is perfect. Once it is perfect, they maintain that reputation. However if their reputation starts below this level, they will not engage in any reputation building, but rather consume what reputation they have. The movements through the state space induced by this optimal strategy are shown in Figure 2.

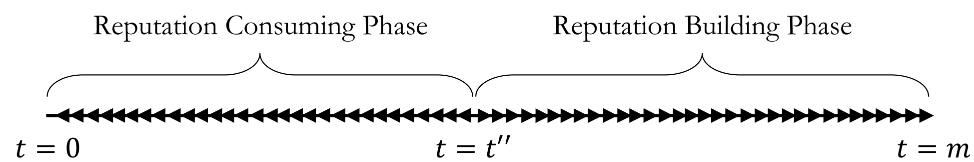

Figure 2: The movements through the state space resulting from the optimal strategy.

Over the long run of the game, where player 1 ends up in the state space depends very much on where they started. If they start below t'', they will end up at $t = 0$ as they will simply consume what reputation they have and never build their reputation. If they start at t'' or above, they will build a perfect reputation. On this basis, it would seem that there is something to be said for giving people 'the benefit of the doubt' when there is no information to proceed on. Doing so may well encourage reputation building behavior to the benefit of society as a whole.

Further insight into the dynamics of the optimal strategy can be found by considering how the critical values, δ', δ'', and t''_{-}/m change as the other parameters of the game change.

$$\frac{d\delta'}{dm} = \frac{(a-b)(b-c)}{m^2\left[(a-b)+\frac{1}{m}(b-c)\right]^2} > 0, \tag{27}$$

$$\frac{d\delta''}{dm} = \frac{(c-d)(b-c)}{m^2\left[(c-d)+\frac{1}{m}(b-c)\right]^2} > 0, \tag{28}$$

$$\frac{\partial\frac{t''_{-}}{m}}{\partial\delta} = -\frac{b-c}{m(1-\delta)^2\left[(c-d)-(a-b)\right]} < 0, \tag{29}$$

$$\frac{\partial\frac{t''_{-}}{m}}{\partial m} = \frac{\frac{\delta}{m^2}(b-c)}{(1-\delta)\left[(c-d)-(a-b)\right]} > 0. \tag{30}$$

The results in (27) and (28) show that the cut off for any reputation building, and the cut off for full reputation building are increasing as the size of the social memory, m increases. This is exactly what we should expect given the effect of the coarsening of the social memory on the costs and benefits of building a reputation.

The result in (29) is extremely intuitive. As player 1 becomes more patient, the cut off between the reputation building and reputation consuming phases in the state space falls. Since the reputation building phase is at the top of the state space, this means that the reputation building phase is expanding at the expense of the reputation consuming phase, which is to be expected as player 1 becomes more patient. Similarly, the result in (30) indicates that as m gets larger, so building a good reputation becomes more expensive, the cut off between the reputation consuming and reputation building phases of the state space will move up the state space. Given that the reputation building phase exists above this cut off, this amounts to an intuitive expansion of the reputation consuming phase at the expense of the reputation building phase.

There is an interesting contrast to draw between the cases of an increasing and decreasing temptation to cheat. Where there is an increasing temptation to cheat, there is the possibility, given an intermediate level of patience, of gravitating to a point in the middle of the state space. No such possibility exists if the temptation to cheat is decreasing. In that scenario, player 1's optimal strategy must take them to one end or the other of the state space. If the player has an intermediate level of patience, then which end of the state space will depend on where in the state space they start.

There is a third possibility which would involve additive separability of player 1's payoffs. Additive separability of player 1's payoffs would mean that $a - b = c - d$ so that the temptation to cheat on a reputation is neither increasing nor decreasing in the probability of being trusted. This essentially makes each state the same as every other state from a strategic perspective, so if it is optimal to build a reputation in one state, it is optimal to build a reputation in all states. This leads to a critical discount factor, δ_s, such that T is the optimal strategy in all states where $\delta \geq \delta_s$, and B is the optimal strategy in all states where $\delta < \delta_s$. The formal derivation of this case has been confined to Appendix C.

4 Comparison with Other Models

Before examining the case where reputations are managed by public relations experts, it is worth pausing to consider some of the insights from the model of unmanaged reputation presented above. A key insight concerns how long lived players might maintain a reputation which is imperfect. If there is an increasing temptation to cheat, and player 1 has an intermediate level of patience, such that $\delta \in [\delta^*, \delta^{**}]$, then player 1's strategy will eventually see them oscillating between the states t' and $t' - 1$. Since $t' \leq m$, the short lived player called to play in a given period has a positive probability of sampling an instance of B, and so not trusting player 1 and playing R. Such a situation is referred to as "reputational failure".

So far as the author is aware, there are no existing reputation models with such a result occurring once the long lived player has built their reputation to the point that they desire. The long lived player may sometimes fail to win trust in the process of building their reputation, but they will normally be trusted once they have built it. In this model, that is not the case, and the reputations that are built may well be less than perfect.

Yet the reputations that are built in the real world are far from perfect. Pick any of the big brands of the day and less than five minutes Googling will suffice to find a bad review of their products or business practices (including Google). Standard models might be able to explain such anomalies as noisy signals resulting from the production process. However this model includes them as an active choice made by the long lived player in a reputation game. Such a possibility should not be excluded. Indeed, the desire for such an imperfect reputation is captured rather nicely by George Orwell: "On the whole human beings want to be good, but not too good, and not quite all the time."[7]

In terms of the assumptions underlying the models the key difference concerns the sophistication of the short lived players. The models of reputation based on incomplete information assume short lived players to know a large amount about the history of their opponent; the possible types of long lived players and be able to use that information to conduct some very sophisticated Bayesian inferences. The starting point for this model is that, even if short lived players were capable of gathering and analyzing information in this way, it would be strange to deploy such impressive and doubtless costly resources for an interaction in which they will engage once.

The actual situation of short lived players in the standard reputation setting is remarkably similar to that of adaptive players in evolutionary game theory. In both cases, they are going to play in an interaction once, or very infrequently, and they have no idea about the motives or incentives of their opponent. But what they might reasonably know about are some of the actions taken by the person playing in the position of their opponent in the past. By

[7] *All Art is Propaganda: Critical Essays*, by George Orwell.

contrast, the long lived player building a reputation has a very large stake in this interaction and so should be expected to commit more cognitive resources to understanding it. Previous papers (Ellison, 1997; Sáez-Martı and Weibull, 1999; Jehiel and Samuelson, 2012) have considered some aspects of how evolutionary players and standard *homo economicus* might interact. However the only one of these to consider reputation issues considered only the fictitious play model, and only in passing. None of these papers dealt with adaptive players in an reputation setting.

Some work (Liu, 2011; Liu and Skrzypacz, 2014) has considered the impact of bounding the memory of the short lived players so that inferences have to be based on only the recent history. This generally results in reputation bubbles where reputations are built up and then 'spent' and then built up again. In these models, lengthening the memory of the short lived players will generally increase the amount of time long lived players spend building their reputations, and reduce the amount of time spent consuming out of them. In the model built here, because a longer memory reduces the marginal impact of any particular act of reputation building or expenditure, longer memories result in worse reputations. In particular, when there is an increasing temptation to cheat and player 1 has an intermediate level of patience, it results in more time spent consuming out of worse reputations and less time spent building them up.

5 Managed Reputation

We now introduce the possibility of hiring a Public Relations (PR) specialist. The PR specialist is able to manipulate the sampling of the short lived player called to move at any point and ensure that, if a T is present in the social memory, then it will be sampled. This might be because they run a strong advertising campaign on the basis of that past incident. Such a campaign is referred to in the public relations field as a 'poster-child campaign'.[8]

Intuitively, the effect of a public relations specialist who can manipulate sampling in this way is that player 1 will be able to maintain a reputation for playing T without actually having to play T so much as in the case where reputations were not managed. Indeed, Proposition 9 shows that, with the

[8]The phrase derives from advertising campaigns used by charities funding medical research into treatments for specific diseases. They typically base these campaigns around a poster featuring a particularly sympathetic looking individual suffering from the disease, normally a child. Hence the name.

help of a public relations specialist, the long lived player will never play T in any state where the number of such actions in the social memory is $t \geq 2$. The argument is relatively simple.

Proposition 9 *When employing the services of a public relations specialist, the long lived player will never play T in any state where $t \geq 2$.*

Proof. Suppose that a candidate strategy for player 1 called for the play of T in a state where $t \geq 2$. Then there would be a profitable single period deviation. Player 1 could play B instead in this state, increasing their payoff by $a - b$. The cost of doing so is that with probability t/m, they will, as a result, move down a state. However, because $t = 1$ is sufficient to be trusted with probability 1, there is no cost to moving down a state in this way. ■

There are two cases, depending on whether (31) holds or not.

$$m > \frac{(b-d)(c-d)}{(a-b)(a-c)}. \tag{31}$$

The case dealt with here is where (31) holds. The case where it does not hold involves some tedious technical issues and is dealt with in Appendix F. Note that (31) must include the case where there is an increasing temptation to cheat. This is because the increasing temptation to cheat means that $a-b > c-d \Rightarrow a-c > b-d$, and so must require that $(b-d)(c-d) / [(a-b)(a-c)] < 1$. So all possible values of m would be greater than this fraction. In fact, since we have assumed that $m > 3$, (31) would also hold over a substantial range of parameter values where there was a decreasing temptation to cheat. In fact the temptation to cheat would have to be decreasing rapidly in order for (31) not to hold for any reasonable values of m.

Then Proposition 10 sets out the optimal strategy for player 1 under various levels of patience.

Proposition 10 *Where the social memory is long enough relative to the payoffs,*

$$m > \frac{(b-d)(c-d)}{(a-b)(a-c)}, \tag{31}$$

the optimal strategy of the long lived player depends on the level of their discount factor, δ as follows:

- $\forall \delta < \delta^{\dagger}$, *play B* $\forall t$;
- $\forall \delta > \delta^{\dagger\dagger}$, *play T for* $t = 0, 1$ *and play B* $\forall t \geq 2$; *and*

- $\forall\delta \in \left[\delta^{\dagger}, \delta^{\dagger\dagger}\right]$, *play T for* $t = 0$ *and play B for* $t \geq 1$.

Where:

$$\delta^{\dagger} = \frac{c-d}{(a-d) - \frac{1}{m}(c-d)}, \tag{32}$$

$$\delta^{\dagger\dagger} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b) + \frac{1}{m}(b-d)}. \tag{33}$$

The proof has been confined to Appendix E, but works in a similar manner to the other proofs. The value function of the proposed equilibrium strategy is calculated and the conditions under which there is no profitable deviation are derived.[9]

An interesting feature of Proposition 10 is that $\delta^{\dagger\dagger} = \delta^{**}$.[10] So the same level of patience is required to lead to the maximal reputation building, whether reputation is managed or not. This is because the trade off between playing T or B at $t = 1$, when reputation is managed, and $t = m$ when reputation is unmanaged is the same. In both cases, the benefit of playing B instead of T is the difference in instantaneous payoff, $a - b$. When reputation is unmanaged, the cost of playing B instead of t is that player 1 will move down to a lower state with certainty, where they will not be trusted with probability $1/m$. When reputation is managed, the cost of playing B in state $t = 1$ is that with probability $1/m$, they will move down to state $t = 0$ and they will not be trusted with certainty. It is not difficult to understand intuitively why the expected cost in terms of foregone future payoffs is the same in both cases.

Another interesting feature of Proposition 10 is that $\delta^{\dagger}$ is actually decreasing in m.

$$\frac{d\delta^{\dagger}}{dm} = -\frac{(c-d)^2}{m^2\left[(a-d) - \frac{1}{m}(c-d)\right]^2} < 0. \tag{34}$$

When in state $t = 1$, as m increases, the probability of moving down from state $t = 1$ to state $t = 2$ as a result of playing B decreases. It just becomes

[9]The case where 31) does not hold has a similar relationship to the case where it does as the relationship between the case of a decreasing and an increasing temptation to cheat when reputation is unmanaged. To be specific, the general pattern of three regions of patience levels for player 1 still pertains when (31) does not hold. At the lowest levels of patience, player 1 never plays T; at the highest levels of patience, player 1 plays T in two states; and at intermediate levels of patience, player 1 plays T in one state. The borders of these levels of patience are different when (31) does not hold. Furthermore, when (31) does not hold, in the intermediate range of discount factors, player 1 will play T at $t = 1$ and B at $t = 0$, rather than T at $t = 0$ and B at $t = 1$ as is the case when (31) holds.

[10]$\delta^{\dagger} < \delta^{*}$ can be shown to be the case $\forall m > 1$.

less likely that the instance of T being used by the public relations expert will be selected to drop out of the social memory. So being in state $t = 1$ becomes more attractive and so less patient players are attracted to playing T instead of B in state $t = 0$ in order to move up to state $t = 1$.

5.1 Comparison of Managed and Unmanaged Reputations

We now compare the results for a firm with a managed reputation to those for a firm with an unmanaged reputation. The focus of the analysis that follows will be the case the temptation to cheat is increasing. This guarantees (31) holds for any m, and so allows for ease of comparison between the two cases, but it also accords quite naturally with most settings. When we trust other people the most is when we are at our most vulnerable to exploitation.

The point of comparison between the two cases will be the relative success of each system in guaranteeing good behavior by player 1, defined by the play of T. The effect of public relations on the behavior of player 1 will depend on that player's level of patience, δ.

So first, where $\delta < \delta^{\dagger}$, the availability of public relations makes no difference to the actions of player 1 and they will always play B. In the long run, they will spend all their time in the state $t = 0$.

However in the case where $\delta \in \left[\delta^{\dagger}, \delta^{*}\right)$ player 1, if they employ a public relations specialist will play T in state $t = 0$, and B in all other states. Whereas a player 1 who does not employ a public relations specialist will continue to play B in all states. So in this situation, a player 1 whose reputation is managed by a PR specialist will play T more frequently than one whose reputation is unmanaged. This is because of the way in which the PR specialist effectively magnifies the benefits of building small reputations.

Now consider the case where $\delta \in [\delta^{*}, \delta_p)$, where δ_p has been defined so that $\delta = \delta_p \Rightarrow t'_{-} = 1$.

$$\delta_p = \frac{m\,(c-d) + [(a-b)-(c-d)]}{(a-c) + \left(1 - \frac{2}{m}\right)[(a-b)-(c-d)] + (m-1)\,(c-d)}. \qquad (35)$$

It follows from $\delta \in [\delta^{*}, \delta_p)$ that $t'_{-} \in [0, 1)$. Since t'_{-} is the lower bound on the bottom of the reputation consuming phase of the state space, this would imply $t' = 1$, so that player 1 plays T in $t = 0$ but B in all other states when their reputation is unmanaged. So the behavior of player 1 is the same whether their reputation is managed or not. However, we should note that player 1 would still benefit from the use of a PR specialist, as they would be more frequently trusted as a result of the specialist's efforts.

Finally, in all cases where $\delta > \delta_p$, player 1 would play T more if their reputation is unmanaged than they would if their reputation were managed. This is because $\delta > \delta_p$ is sufficient to ensure that they play T at $t = 1$ as well as $t = 0$ when reputations are unmanaged. To play T at $t = 1$ when reputation is managed required $\delta > \delta^{**}$, but a player this patient would play T in all states if reputation were unmanaged.

These situations are shown in Figure 3. It should be clear, since $d\delta^{\dagger}/dm < 0$, and $d\delta^{*}/dm > 0$, that as m increases, δ^{*} increases and $\delta^{\dagger}$ decreases so that the gap between them widens. This means the parameter space where player 1 will build a play T more often if their reputation is managed increases. It can similarly be shown that δ_p is increasing in m. So the region of parameter space where player 1 will play T more frequently if their reputation is unmanaged than if it is managed is shrinking as m increases. But the gap between δ^{*} and δ_p can be shown to shrink as m increases. This leads to the intuitive conclusion that as memories get longer, public relations to manipulate how those memories are used becomes not only more valuable, but, potentially, of greater social utility as well.

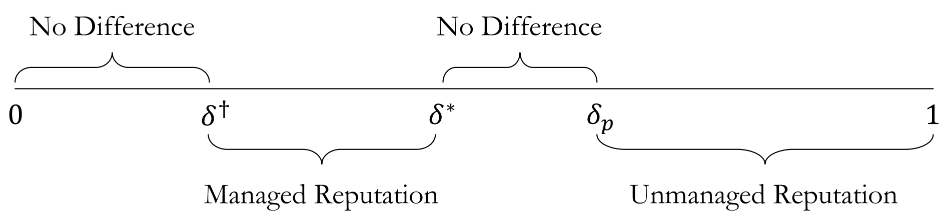

Figure 3: The levels of patience which imply players with managed reputations behave better.

So whether the availability of public relations improves the average behavior of player 1 depends on the player's patience. For very impatient players, $\delta < \delta^{\dagger}$ there is no difference. However, there are levels of patience, $\delta \in \left[\delta^{\dagger}, \delta^{*}\right)$ where player 1 will play T more frequently as a result of the way in which a PR specialist magnifies the benefit of even small reputations. There is then a range of parameters $\delta \in [\delta^{*}, \delta_p)$ where the behavior of player 1 will be the same whether they have the benefit of a PR specialist's efforts or not. Finally, $\forall \delta \geq \delta_p$, a player 1 whose reputation is unmanaged behaves better than one whose reputation is managed. This better behavior comes about because of the way in which a PR specialist magnifies the reputation benefit of being in low states. As a result, a player 1 using their services does not need to behave any better.

6 Conclusions and Avenues for Future Research

This paper departs from the incomplete information setting of reputations which has been standard since the 1980s. Instead, given the similarity between short lived players in a reputation game and adaptive players in evolutionary game theory, the model of adaptive play has been adapted and applied in a reputation setting. The long lived player building a reputation remains a standard *homo economicus* and exploits the adaptive expectations of the short lived player.

The extent to which the long lived player constructs a reputation depends on how patient they are. If they exhibit intermediate levels of patience, they might optimally accept some positive probability that their reputation is insufficient to ensure that they are always trusted. This produces intuitive reputation cycling effects through a much simpler model than previous papers.

Furthermore, the sampling behavior of adaptive players provides an intuitive way to model the actions of Public Relations specialists in terms of manipulating the sample drawn by short lived players. The results are interesting in that at low levels of patience, PR strategists might encourage individuals who would not otherwise bother to build some reputation. In doing so, they increase the incidence of "good behavior". However at higher levels of patience, PR specialists increase the amount of cheating on a reputation that takes place. Why build a golden reputation when a gold plated one will be sufficient to ensure trust.

The key issue in modeling the effects of a PR specialist is not the adaptive behavior of the short lived players, but the fact that they are drawing inferences on the basis of a sample that could be manipulated. It might well be possible to model PR specialists in a standard reputation model, but where Bayesian inferences are based on a sampling of recent past actions. However, the simpler model shown here is sufficient for the purpose of examining the impact of PR specialists, so why build a more complicated one?

Furthermore, this way of thinking about reputations by reducing the rationality of the short lived players, opens up some extremely interesting avenues for future research. First, it would be interesting to see how a similar model might be applied to a class of games known as audit games, where the short lived players are deciding whether to cheat or not and the long lived player is deciding whether to engage in a costly audit which will only reap benefits if the short lived player has actually cheated on that occasion. The impact of public relations specialists in this setting will be of particular interest to residents of the United Kingdom where the BBC frequently engages in costly advertising to claim an ability to catch those who don't pay the annual TV License Fee.[11]

A second area of further research involves moving beyond a sample size of $s = 1$. Allowing short lived players to draw larger samples would add further realism to the model, at the cost of greater complexity. However the insights gained from such an addition to the model may well justify the greater complexity.

A third avenue for future research would involve considering how reputations are built when the stage game will be repeated only a finite number of times. The effect in terms of the model will be to change the nature of the dynamic programming problem for player 1. It would be intuitive to expect that player 1, if the game is long enough, would start by building up their reputation, but that the reputation spending region of the state space would be expected to expand as the game neared its end point.

A model of reputation which does not use incomplete information may now seem somewhat unconventional. Nevertheless, using the developments from evolutionary game theory in recent years to probe how reputations work when short lived players do not bring their fully calculating faculties to bear offers some useful insights. In particular, we see reputations which are built, but are not built to their fullest potential and are not perfectly maintained. The usual assumption of fully rational short lived players stresses the fact that it is impossible to fool all of the people all of the time. Reduced rationality for short lived players, along the lines of those that are used in evolutionary game theory offer a timely reminder that it remains possible to fool some of the people some of the time.

[11] The TV license fee is the annual charge levied on all households with a Television in order to fund public service broadcasters.

References

Aperjis, Christina, Richard J. Zeckhauser, and Yali Miao. 2014. "Variable temptations and black mark reputations." *Games and Economic Behavior*, 87: 70 – 90.

Bagwell, Kyle. 2007. "Chapter 28 The Economic Analysis of Advertising." In . Vol. 3 of *Handbook of Industrial Organization*, , ed. M. Armstrong and R. Porter, 1701 – 1844. Elsevier.

Cripps, Martin W., George J. Mailath, and Larry Samuelson. 2004. "Imperfect Monitoring and Impermanent Reputations." *Econometrica*, 72(2): 407–432.

Ekmekci, Mehmet, and Andrea Wilson. 2007. "Reputation Games with Finite Automata." mimeo.

Ellison, Glenn. 1997. "Learning from Personal Experience: One Rational Guy and the Justification of Myopia." *Games and Economic Behavior*, 19(2): 180 – 210.

Fudenberg, Drew, and David K Levine. 1989. "Reputation and Equilibrium Selection in Games with a Patient Player." *Econometrica*, 57(4): 759–78.

Fudenberg, Drew, and David K. Levine. 1992. "Maintaining a Reputation when Strategies are Imperfectly Observed." *The Review of Economic Studies*, 59(3): 561–579.

Jehiel, P. 2005. "Analogy-based expectation equilibrium." *Journal of Economic theory*, 123(2): 81–104.

Jehiel, Philippe, and Larry Samuelson. 2012. "Reputation with Analogical Reasoning*." *The Quarterly Journal of Economics*, 127(4): 1927–1969.

Kandori, Michihiro, George J. Mailath, and Rafael Rob. 1993. "Learning, Mutation, and Long Run Equilibria in Games." *Econometrica*, 61(1): pp. 29–56.

Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson. 1982. "Rational cooperation in the finitely repeated prisoners' dilemma." *Journal of Economic Theory*, 27(2): 245–252.

Kreps, D.M., and R. Wilson. 1982. "Reputation and imperfect information." *Journal of economic theory*, 27(2): 253–279.

Liu, Qingmin. 2011. "Information acquisition and reputation dynamics." *The Review of Economic Studies*, 78(4): 1400–1425.

Liu, Qingmin, and Andrzej Skrzypacz. 2014. "Limited records and reputation bubbles." *Journal of Economic Theory*, 151: 2–29.

Mailath, George J., and Larry Samuelson. 2001. "Who Wants a Good Reputation?" *The Review of Economic Studies*, 68(2): pp. 415–441.

Milgrom, Paul, and John Roberts. 1982. "Predation, reputation, and entry deterrence." *Journal of Economic Theory*, 27(2): 280 – 312.

Nelson, Phillip. 1974. "Advertising as Information." *Journal of Political Economy*, 82(4): pp. 729–754.

Phelan, Christopher. 2006. "Public trust and government betrayal." *Journal of Economic Theory*, 130(1): 27 – 43.

Sáez-Martı, Marıa, and Jörgen W Weibull. 1999. "Clever agents in Young's evolutionary bargaining model." *Journal of Economic Theory*, 86(2): 268–279.

Selten, Reinhard. 1978. "The chain store paradox." *Theory and Decision*, 9: 127–159. 10.1007/BF00131770.

Wiseman, Thomas. 2008. "Reputation and impermanent types." *Games and Economic Behavior*, 62(1): 190 – 210.

Young, H. Peyton. 1993. "The Evolution of Conventions." *Econometrica*, 61(1): 57–84.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton, NJ:Princeton University Press.

A Proofs of Lemma 1 & 3

This Appendix gives the proofs for Lemmas 1 and 3, which are very similar.

A.1 Proof of Lemma 1

Lemma 1 *When player 1 always plays B between state* 0 *and state* t^*, *the change in the value function,* $\mathbf{D}V(t)$ *does not depend on the state,* t $\forall t \in \{1, 2, \ldots t^*\}$, *and for all these values is equal to:*

$$\mathbf{D}V(t) = \frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m} \tag{9}$$

Proof. The proof works by induction. First, we will show that $\mathbf{D}V(1) = \mathbf{D}V(2)$. We then show tjhat if $\mathbf{D}V(t-1) = \mathbf{D}V(t-2)$, then it must be the case that $\mathbf{D}V(t) = \mathbf{D}V(t-1)$.

For the first part, note that

$$V(0) = \frac{c}{1-\delta}. \tag{36}$$

By the rearranging the value function, we can also find $V(1)$ and $V(2)$:

$$V(1) = \frac{(1-\delta+\delta/m)\,c + \frac{1}{m}(1-\delta)(a-c)}{(1-\delta)(1-\delta+\delta/m)}.$$

$$V(2) = \frac{(1-\delta+\delta/m)\,c + 2(1-\delta)\frac{1}{m}(a-c)}{(1-\delta)(1-\delta+\delta/m)}.$$

Hence:

$$V(1) - V(0) = \frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m}. \tag{37}$$

$$V(2) - V(1) = \frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m} \tag{38}$$

And so $\mathbf{D}V(2) = \mathbf{D}V(1)$.

Now, an implication of (8) is:

$$\mathbf{D}V(t) - \mathbf{D}V(t-1) = \frac{\frac{1}{m}(a-c) - (1-\delta+\delta/m)\,\mathbf{D}V(t-1)}{1-\delta\left(1-\frac{t}{m}\right)} \tag{39}$$

So the growth in $\mathbf{D}V(t)$ will be equal to zero where

$$\mathbf{D}V(t-1) = \frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m}. \tag{40}$$

This was the case at $\mathbf{D}V(1)$, and so must also have been the case at $\mathbf{D}V(2)$ and $\mathbf{D}V(3)$ and so on for all values of t between 1 and t^*. ■

A.2 Proof of Lemma 3

Lemma 3 *When player 1 always plays T between state t^* and state m, the change in the value function, $\mathbf{D}V(t)$ does not depend on the state, t $\forall t \in \{t^*+1, t^*+2, \ldots m\}$, and for all these values is equal to:*

$$\mathbf{D}V(t) = \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m} \tag{13}$$

Proof. The proof once again works by induction. We first show that

$$\mathbf{D}V(m) = \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m}. \tag{41}$$

We then show that in general, if

$$\mathbf{D}V(t+1) = \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m},$$

then $\mathbf{D}V(t) = \mathbf{D}V(t+1)$, and so by induction, $\mathbf{D}V(t) = \mathbf{D}V(t+1)$ $\forall t \in \{t^*, t^*+1, m-1\}$.

Clearly,

$$V(m) = \frac{b}{1-\delta}. \tag{42}$$

Then

$$V(m-1) = \frac{(1-\delta+\delta/m)\,b - (1-\delta)\frac{1}{m}(b-d)}{(1-\delta)(1-\delta+\delta/m)}. \tag{43}$$

Therefore:

$$\mathbf{D}V(m) = \frac{b}{1-\delta} - \frac{(1-\delta+\delta/m)\,b-(1-\delta)\frac{1}{m}(b-d)}{(1-\delta)(1-\delta+\delta/m)} = \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m}. \tag{41}$$

We know from (12) that in general:

$$\mathbf{D}V(t) - \mathbf{D}V(t+1) = \frac{\frac{1}{m}(b-d) - (1-\delta+\delta/m)\,\mathbf{D}V(t+1)}{1-\delta\frac{t-1}{m}}. \tag{44}$$

And this shows that:

$$\mathbf{D}V(t) = \mathbf{D}V(t+1) \Leftrightarrow \mathbf{D}V(t+1) = \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m}. \tag{45}$$

Hence $\mathbf{D}V(t^*+1) = \mathbf{D}V(t^*+2) = \ldots = \mathbf{D}V(m-1) = \mathbf{D}V(m)$. As required. ■

B Proof of Proposition 5

Proposition 5 *If $\delta \in [\delta^*, \delta^{**}]$, and $a - b > c - d$, then the optimal strategy for player 1 involves building a reputation and playing T when $t < t'$ and consuming out of their reputation and playing B where $t \geq t'$, where t' is the sole integer in the interval $\left[t'_-, t'_- + 1\right]$, and*

$$t'_- = \frac{\delta (a - c) - m (1 - \delta + \delta/m) (c - d)}{(1 - \delta + 2\delta/m) [(a - b) - (c - d)]} \tag{16}$$

Several key Lemmas must be stated and proved before the main proposition can be proved, these Lemmas are stated and proved in the following subsections. Once they have been established, Proposition, 5 will follow quite easily.

B.1 Establishing t'_-

Lemma 6 *If player 1's optimal strategy is to play T $\forall t < t'$ and to play T $\forall t \geq t'$, then the only possible value of t' is the sole integer in the range $\left(t'_-, t'_- + 1\right)$, where*

$$t'_- = \frac{\delta (a - c) - m (1 - \delta + \delta/m) (c - d)}{(1 - \delta + 2\delta/m) [(a - b) - (c - d)]} \tag{16}$$

Furthermore, if $\delta \in [\delta^*, \delta^{**}]$, then $t'_- \in [0, m]$.

Proof. The change in the value function moving from $t' - 1$ to t' can be defined without reference to any level of or change in the value function at any other point of the state space.

$$\mathbf{D}V\left(t'\right) = \frac{\frac{t'}{m}(a - b) + \left(1 - \frac{t'}{m}\right)(c - d) + \frac{1}{m}(b - d)}{1 + \delta/m}. \tag{46}$$

The differences in the value function in consecutive states where the player is playing T and where the player is playing B have already been calculated. So a summary of all the possible values for the difference in the value function under this candidate strategy is:

$$\mathbf{D}V(t) = \begin{cases} \frac{\frac{1}{m}(b-d) + \delta\left(1 - \frac{t}{m}\right)\mathbf{D}V(t+1)}{1 - \delta\frac{t-1}{m}} & \forall t < t' \\ \frac{\frac{t'}{m}(a-b) + \left(1 - \frac{t'}{m}\right)(c-d) + \frac{1}{m}(b-d)}{1 + \delta/m} & t = t' \\ \frac{\frac{1}{m}(a-c) + \delta\frac{t-1}{m}\mathbf{D}V(t-1)}{1 - \delta\left(1 - \frac{t}{m}\right)} & \forall t > t' \end{cases}. \tag{47}$$

Using (47), expressions for $\mathbf{D}V(t'-1)$ and $\mathbf{D}V(t'+1)$ can be found. For now these are left in terms of $\mathbf{D}V(t')$.

$$\mathbf{D}V\left(t'-1\right) = \frac{\frac{1}{m}(b-d)}{1-\delta\frac{t'-2}{m}} + \frac{\delta\left(1-\frac{t'-1}{m}\right)\mathbf{D}V(t')}{1-\delta\frac{t'-2}{m}}, \tag{48}$$

$$\mathbf{D}V\left(t'+1\right) = \frac{\frac{1}{m}(a-c)}{1-\delta\left(1-\frac{t'+1}{m}\right)} + \frac{\delta\frac{t'}{m}\mathbf{D}V(t')}{1-\delta\left(1-\frac{t'+1}{m}\right)}. \tag{49}$$

These changes in the value function can be substituted into the $H(t)$ function at t'

$$\begin{aligned} H\left(t'\right) &= \delta\frac{t'}{m}\mathbf{D}V\left(t'\right) + \delta\left(1-\frac{t'}{m}\right)\mathbf{D}V\left(t'+1\right), \\ &= \frac{\frac{\delta}{m}\left(1-\frac{t'}{m}\right)(a-c) + \delta\frac{t'}{m}(1+\delta/m)\mathbf{D}V(t')}{1-\delta\left(1-\frac{t'+1}{m}\right)}. \end{aligned} \tag{50}$$

So the requirement that $H(t') < G(t')$ means that:

$$\begin{aligned} &\frac{\delta}{m}\left(1-\frac{t'}{m}\right)(a-c) + \frac{\delta}{m}\frac{t'}{m}(b-d) \\ &\qquad < (1-\delta+\delta/m)\left[(c-d) + \frac{t'}{m}\left[(a-b)-(c-d)\right]\right]. \end{aligned}$$

Which can be rearranged to:

$$\frac{t'}{m} > \frac{t'_-}{m} = \frac{\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)\left[(a-b)-(c-d)\right]} \tag{51}$$

Next, substitute the relevant changes in the value function into $H(t)$ at $t'-1$:

$$\begin{aligned} H\left(t'-1\right) &= \delta\frac{t'-1}{m}\mathbf{D}V\left(t'-1\right) + \delta\left(1-\frac{t'-1}{m}\right)\mathbf{D}V\left(t'\right), \\ &= \frac{\frac{\delta}{m}\frac{t'-1}{m}(b-d) + \delta\left(1-\frac{t'-1}{m}\right)(1+\delta/m)\mathbf{D}V(t')}{1-\delta\frac{t'-2}{m}}. \end{aligned}$$

So the requirement that $H(t'-1) > G(t'-1)$ can be written as

$$\begin{aligned} &\frac{\delta}{m}\frac{t'-1}{m}(b-d) + \frac{\delta}{m}\left(1-\frac{t'-1}{m}\right)(a-c) \\ &\qquad > (1-\delta+\delta/m)\left[(c-d) + \frac{t'-1}{m}\left[(a-b)-(c-d)\right]\right], \end{aligned}$$

The simplest approach from here is to arrange as a condition on $(t'-1)/m$:

$$\frac{t'-1}{m} < \frac{t'_- - 1}{m} = \frac{\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)[(a-b)-(c-d)]}. \tag{52}$$

Taking (51) and (52) together gives the requirement that $t' \in \left(t'_-, t'_- + 1\right)$, where

$$t'_- = \frac{\delta(a-c) - m(1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)[(a-b)-(c-d)]}. \tag{16}$$

Finally check that t' is actually in the available state space. This will require that $t'_- \in [0, m]$. The requirement of $t'_- > 0$ gives:

$$\begin{aligned} &\frac{\delta}{m}(a-c) > (1-\delta+\delta/m)(c-d), \\ &\Rightarrow \delta > \delta^* = \frac{c-d}{\left(1-\frac{1}{m}\right)(c-d) + \frac{1}{m}(a-c)}. \end{aligned} \tag{53}$$

The requirement that $t'_- < m$, gives:

$$\begin{aligned} \frac{\delta}{m}(a-c) - (1-\delta+\delta/m) &< (1-\delta+2\delta/m)[(a-b)-(c-d)], \\ \Rightarrow \delta < \delta^{**} &= \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b) + \frac{1}{m}(b-d)}. \end{aligned} \tag{54}$$

The requirements of (53) and (54) together give the overall requirement that $\delta \in [\delta^*, \delta^{**}]$. ■

This argument effectively establishes t'_- by finding limits on the crossing point between $G(t)$ and $H(t)$. The next step is to establish that this crossing point must be unique. This will involve proving that it must be optimal for player 1 to play T whenever $t \le t'-2$, and to play B whenever $t \ge t'+1$. The strategy here is to show that $H(t)$ under the proposed strategy is an increasing function. For $t > t'$, the rate of increase is falling, and the initial rate of increase is less than the rate of increase in $G(t)$. This makes it impossible, for $t > t'$, for $H(t)$ to overtake $G(t)$, and so it must always be the case that B is the optimal action.

However, for $t < t'-1$, the rate at which $H(t)$ increases is itself increasing. Greater clarity can be obtained by considering the consequences of this from the perspective of the state, t'. It means that as the state moves away from t' towards lower states, the function $H(t)$ decreases and at a decreasing rate (the absolute value of the decrease falls). Meanwhile, $G(t)$ decreases at a constant rate. By definition, $G(t)$ was below $H(t)$ at $t'-1$. If the decrease in $H(t)$ between $t'-2$ and $t'-1$ was not enough to bring the function $H(t)$ below $G(t)$, none of the subsequent decreases will be sufficient.

In finding the properties of the function $H(t)$, the shape of the $\mathbf{D}V(t)$ function are crucial.

B.2 The Properties of $\mathbf{D}V(t)$

Lemma 7 *Under the candidate strategy profile, and where $a-b>c-d$, the function $\mathbf{D}V(t)$ is an increasing function,*

$$\mathbf{D}V(t+1) > \mathbf{D}V(t). \tag{55}$$

For all $t<t'$, the rate of increase is increasing,

$$\mathbf{D}V(t+2) - \mathbf{D}V(t+1) > \mathbf{D}V(t+1) - \mathbf{D}V(t), \tag{56}$$

For all $t>t'$, the rate of increase if decreasing,

$$\mathbf{D}V(t+2) - \mathbf{D}V(t+1) < \mathbf{D}V(t+1) - \mathbf{D}V(t), \tag{57}$$

Proof. The proof is quite lengthy and proceeds in four sections.

1. Showing that $\mathbf{D}V(t)$ is an increasing function for $t>t'$.
2. Showing that $\mathbf{D}V(t)$ is an increasing function for $t<t'$.
3. Showing that $\mathbf{D}V(t)$ increases at a decreasing rate for $t>t'$.
4. Showing that $\mathbf{D}V(t)$ increases at an increasing rate for $t<t'$.

B.2.1 $\mathbf{D}V(t)$ is increasing for $t>t'$

It has already been shown that in a reputation consuming phase:

$$\mathbf{D}V(t) - \mathbf{D}V(t-1) = \frac{\frac{1}{m}(a-c) - (1-\delta+\delta/m)\mathbf{D}V(t-1)}{1-\delta\left(1-\frac{t}{m}\right)} \tag{39}$$

So $\mathbf{D}V(t) - \mathbf{D}V(t-1) > 0$, if and only if

$$\mathbf{D}V(t-1) < \frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m}. \tag{58}$$

Applying the formula for $\mathbf{D}V(t-1)$ from (8) gives:

$$\frac{\frac{1}{m}(a-c)+\delta\frac{t-2}{m}\mathbf{D}V(t-2)}{1-\delta\left(1-\frac{t-1}{m}\right)}<\frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m},$$
$$\Rightarrow \mathbf{D}V(t-2)<\frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m}. \tag{59}$$

Obviously, (59) is the condition for $\mathbf{D}V(t-1)-\mathbf{D}V(t-2)>0$. So $\mathbf{D}V(t)>\mathbf{D}V(t-1)\Leftrightarrow \mathbf{D}V(t-1)>\mathbf{D}V(t-2)$. This logic will extend all the way back to the state $t'+1$.

Next, consider:

$$\mathbf{D}V\left(t'+1\right)-\mathbf{D}V\left(t'\right)=\frac{\frac{1}{m}(a-c)-(1-\delta+\delta/m)\,\mathbf{D}V\left(t'\right)}{1-\delta\left(1-\frac{t'+1}{m}\right)}. \tag{60}$$

Clearly:

$$\mathbf{D}V\left(t'+1\right)>\mathbf{D}V\left(t'\right)\Leftrightarrow \mathbf{D}V\left(t'\right)<\frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m},$$

Which will be true if and only if

$$\frac{\frac{t'}{m}(a-b)+\left(1-\frac{t'}{m}\right)(c-d)+\frac{1}{m}(b-d)}{1+\delta/m}<\frac{\frac{1}{m}(a-c)}{1-\delta+\delta/m},$$

Which requires:

$$\frac{t'}{m}<\frac{(1+\delta/m)\frac{1}{m}\left[(a-b)-(c-d)\right]+\frac{\delta}{m}(b-d)-(1-\delta+\delta/m)(c-d)}{(1-\delta+\delta/m)\left[(a-b)-(c-d)\right]}. \tag{61}$$

The condition in (61) places a maximum value on t'/m. However, we have already found a range for t'/m, which defines that value. It can be shown that, given $\delta\in(\delta^{*},\delta^{**})$, t'/m must satisfy the condition in (61). This must be true, because within this range of values for δ, the previously established maximum value of t', $\left(t'_{-}+1\right)/m$ must be smaller than the right hand side of (61).

$$\frac{(1+\delta/m)\frac{1}{m}\left[(a-b)-(c-d)\right]+\frac{\delta}{m}(b-d)-(1-\delta+\delta/m)(c-d)}{(1-\delta+\delta/m)\left[(a-b)-(c-d)\right]}$$
$$>\frac{\frac{\delta}{m}(a-c)-(1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)\left[(a-b)-(c-d)\right]}+\frac{1}{m}.$$

Which is equivalent to the following expression:

$$\frac{\delta}{m}\left[(a-c)-(b-d)+(b-d)\right] > (1-\delta+\delta/m)(c-d),$$

$$\Rightarrow \delta > \delta^* = \frac{c-d}{\left(1-\frac{1}{m}\right)(c-d)+\frac{1}{m}(a-c)} \tag{62}$$

So it must be the case, for the parameter values we are considering, that $\mathbf{D}V(t'+1) > \mathbf{D}V(t')$. And so it must also be the case, by induction, that $\mathbf{D}V(t) > \mathbf{D}V(t-1)\,\forall t > t'$. This proves the first part of the lemma, that $\mathbf{D}V(t)$ is an increasing function where $t > t'$.

B.2.2 $\mathbf{D}V(t)$ is increasing for $t < t'$

Now consider the case where $t < t'$, and the candidate strategy calls for the play of T. In a region of the state space where player 1 is playing T,

$$\mathbf{D}V(t) - \mathbf{D}V(t+1) = \frac{\frac{1}{m}(b-d) - (1-\delta+\delta/m)\,\mathbf{D}V(t+1)}{1-\delta\frac{t-1}{m}},$$

For $\mathbf{D}V(t)$ to be an increasing function, we need $\mathbf{D}V(t) - \mathbf{D}V(t+1) < 0$, which requires:

$$\frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m} < \mathbf{D}V(t+1),$$

$$\Rightarrow \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m} < \frac{\frac{1}{m}(b-d)+\delta\left(1-\frac{t+1}{m}\right)\mathbf{D}V(t+2)}{1-\delta\frac{t}{m}},$$

$$\Rightarrow \frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m} < \mathbf{D}V(t+2). \tag{63}$$

Clearly, (63) is also the condition required for $\mathbf{D}V(t+1) - \mathbf{D}V(t+2) < 0$. So

$$\mathbf{D}V(t) - \mathbf{D}V(t+1) < 0 \Leftrightarrow \mathbf{D}V(t+1) - \mathbf{D}V(t+2) < 0.$$

So if it could be shown that $\mathbf{D}V(t'-1) - \mathbf{D}V(t') < 0$, that would mean that $\mathbf{D}V(t) - \mathbf{D}V(t+1) < 0$ for all $t \leq t'-1$, and so prove that the function $\mathbf{D}V(t)$ is increasing over the range we are considering.

$$\mathbf{D}V\left(t'-1\right) - \mathbf{D}V\left(t'\right) = \frac{\frac{1}{m}(b-d) - (1-\delta+\delta/m)\,\mathbf{D}V(t')}{1-\delta\frac{t'-2}{m}},$$

$$\mathbf{D}V\left(t'-1\right) < \mathbf{D}V\left(t'\right) \Rightarrow \frac{1}{m}(b-d) < (1-\delta+\delta/m)\,\mathbf{D}V\left(t'\right).$$

Which becomes:

$$\frac{\frac{1}{m}(b-d)}{1-\delta+\delta/m} < \frac{\frac{t'}{m}(a-b)+\left(1-\frac{t'}{m}\right)(c-d)+\frac{1}{m}(b-d)}{1+\delta/m},$$

$$\Rightarrow \frac{t'}{m} > \frac{\frac{\delta}{m}(b-d)-(1-\delta+\delta/m)(c-d)}{(1-\delta+\delta/m)\left[(a-b)-(c-d)\right]}. \tag{64}$$

Note that such a condition, which places a minimum value on t'/m will be redundant, if the right hand side of (64) is less than the minimum value we know t'/m must take anyway, t'_{-}/m. This requires that

$$\frac{\frac{\delta}{m}(a-c)-(1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)\left[(a-b)-(c-d)\right]} > \frac{\frac{\delta}{m}(b-d)-(1-\delta+\delta/m)(c-d)}{(1-\delta+\delta/m)\left[(a-b)-(c-d)\right]},$$

Cross multiplication and cancellation of terms gives:

$$(1-\delta+\delta/m)\left[(a-b)-(c-d)\right] > \frac{\delta}{m}(b-d)-(1-\delta+\delta/m)(c-d),$$

$$\Rightarrow \delta < \delta^{**} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)}. \tag{65}$$

So in the region we are considering, $\mathbf{D}V(t)$ is an increasing function.

So the first part of the Lemma has been proved and it has been demonstrated that $\mathbf{D}V(t)$ is an increasing function of t. It only remains to demonstrate that it increases at a decreasing rate for $t > t'$, and at an increasing rate for $t < t'$.

B.2.3 $\mathbf{D}V(t)$ is increasing at a decreasing rate for $t > t' + 1$

To prove the next part of the Lemma, consider $\beta = [\mathbf{D}V(t) - \mathbf{D}V(t-1)] - [\mathbf{D}V(t-1) - \mathbf{D}V(t-2)]$, where $t > t' + 1$. Applying the definitions above, this can be written as:

$$\beta = \frac{\frac{1}{m}(a-c)}{1-\delta\left(1-\frac{t}{m}\right)} - \frac{\frac{1}{m}(a-c)}{1-\delta\left(1-\frac{t-1}{m}\right)} - \frac{(1-\delta+\delta/m)\,\mathbf{D}V(t-1)}{1-\delta\left(1-\frac{t}{m}\right)} + \frac{(1-\delta+\delta/m)\,\mathbf{D}V(t-2)}{1-\delta\left(1-\frac{t-1}{m}\right)}.$$

Rearranging and applying the formula for $\mathbf{D}V(t) - \mathbf{D}V(t-1)$ that was developed above, this becomes:

$$\beta = -\frac{\frac{\delta}{m}\left[\mathbf{D}V(t)-\mathbf{D}V(t-1)\right]+(1-\delta+\delta/m)\left[\mathbf{D}V(t-1)-\mathbf{D}V(t-2)\right]}{1-\delta\left(1-\frac{t-1}{m}\right)}.$$

Since all the terms in the fraction are positive, it follows that $\beta < 0$, and so the rate at which the function $\mathbf{D}V(t)$ increases must be decreasing.

B.2.4 $\mathbf{D}V(t)$ is increasing at an increasing rate for $t > t' + 1$

The final part of the Lemma involves demonstrating that $\mathbf{D}V(t)$ is increasing at an increasing rate. For notational convenience, define γ as the difference between the increase in the function $\mathbf{D}V(t)$ from $t+1$ to t and from $t+2$ to $t+1$, so that $\gamma = [\mathbf{D}V(t) - \mathbf{D}V(t+1)] - [\mathbf{D}V(t+1) - \mathbf{D}V(t+2)]$. Then the requirement is $\gamma > 0$.

$$\gamma = \frac{1}{m}(b-d)\left[\frac{1}{1-\delta\frac{t-1}{m}} - \frac{1}{1-\delta\frac{t}{m}}\right] - (1-\delta+\delta/m)\left[\frac{\mathbf{D}V(t+1)}{1-\delta\frac{t-1}{m}} - \frac{\mathbf{D}V(t+2)}{1-\delta\frac{t}{m}}\right]$$

Which becomes:

$$\gamma = -\frac{\frac{\delta}{m}[\mathbf{D}V(t+1) - \mathbf{D}V(t+2)]}{\left(1-\delta\frac{t}{m}\right)\left(1-\delta\frac{t-1}{m}\right)} - \frac{(1-\delta+\delta/m)[\mathbf{D}V(t+1) - \mathbf{D}V(t+2)]}{\left(1-\delta\frac{t}{m}\right)\left(1-\delta\frac{t-1}{m}\right)} > 0. \tag{66}$$

The inequality follows from the fact we know that $\mathbf{D}V(t+1) - \mathbf{D}V(t+2) < 0$, by virtue of the fact that $\mathbf{D}V(t)$ is an increasing function, which was shown in the first part of the lemma.

This competes the proof of the Lemma. ■

B.3 The Implications for $H(t)$

From Lemma 7, Corollary 8 follows.

Corollary 8 *Under the candidate strategy profile, and where $a - b > c - d$, the function $H(t)$ is an increasing function,*

$$H(t+1) > H(t). \tag{67}$$

For all $t < t'$, the rate of increase is increasing:

$$H(t+2) - H(t+1) > H(t+1) - H(t), \tag{68}$$

and for all $t > t'$, the rate of increase is falling,

$$H(t+2) - H(t+1) < H(t+1) - H(t). \tag{69}$$

Proof. Let $\mathbf{D}H(t) = H(t) - H(t-1)$, then

$$\mathbf{D}H(t) = \delta \frac{t-1}{m} [\mathbf{D}V(t) - \mathbf{D}V(t-1)] + \delta \left(1 - \frac{t}{m}\right) [\mathbf{D}V(t+1) - \mathbf{D}V(t)]. \tag{70}$$

Since all the terms in this expression are positive, by Lemma 7, it follows that $\mathbf{D}H(t) > 0$. Furthermore:

$$\begin{aligned} &\mathbf{D}H(t) - \mathbf{D}H(t-1) \\ &\quad = \delta \frac{t-2}{m} [(\mathbf{D}V(t) - \mathbf{D}V(t-1)) - (\mathbf{D}V(t-1) - \mathbf{D}V(t-2))] \\ &\quad + \delta \left(1 - \frac{t}{m}\right) [(\mathbf{D}V(t+2) - \mathbf{D}V(t+1)) - (\mathbf{D}V(t+1) - \mathbf{D}V(t))] \end{aligned} \tag{71}$$

Again, by Lemma 7, we know that all the terms in (71) are negative where $t > t'$ and positive where $t < t'$. So the whole expression must be negative where $t > t'$ and positive where $t < t'$. This proves that $H(t)$ is an increasing function, but increases at a decreasing rate where $t > t'$ and at an increasing rate where $t < t'$.[12] ■

B.4 The Optimality of the Candidate Strategy at all Other States

We know that $G(t)$ is an increasing function which increases at a constant rate in t.

$$\mathbf{D}G(t) = G(t) - G(t-1) = \frac{1}{m} [(a-b) - (c-d)] \tag{72}$$

So it is known that at t', $G(t') > H(t')$. It is also known that for $t > t'$, both $G(t)$ and $H(t)$ will be increasing as t increases. In the case of $G(t)$, the rate of increase is constant. In the case of $H(t)$, the rate of increase is

[12] There might be the technical objection that for $t < 2$, the term in $(t-2)/m$ will be negative and so the preceding reasoning won't follow. However, note that this is essentially a second difference, and second differences as a concept don't make sense for $t < 2$. Similar objections regarding (70) can be met with n equivalent response.

decreasing. So if it can be shown that in the movement from t' to $t'+1$, $G(t)$ grew by more than $H(t)$, then there is no hope of $H(t)$ ever overtaking $G(t)$ between $t = t'$ and $t = m$. So in all these states, it must indeed be optimal to play B rather than T.

Similarly, it is known that at $t'-1$, $G(t'-1) < H(t'-1)$. As the state moves further away from t' to lower values of t, both the $G(t)$ and the $H(t)$ functions will fall. The $G(t)$ function falls at a constant rate, but the $H(t)$ function will fall at a decreasing rate. So if it can be shown that in the move from state $t'-1$ to state $t'-2$, $H(t)$ does not fall below $G(t)$, there will be no chance of $H(t)$ ever falling below $G(t)$ between states $t = t'-1$ and $t = 0$, so T must be the optimal action in all of these states.

Lemma 9 *Given the candidate strategy profile and an increasing temptation to cheat, $a-b > c-d$, the growth of $H(t)$ from $t = t'$ to $t = t'+1$ is less than the growth of $G(t)$ between the same two states; and, provided $\delta \leq \delta^{**}$, the fall in $H(t)$ from $t'-1$ to $t'-2$ is less than the fall in $G(t)$ between the same two states:*

$$\mathbf{D}H\left(t'+1\right) < \mathbf{D}G\left(t'+1\right), \tag{73}$$

$$\mathbf{D}G\left(t'-1\right) > \mathbf{D}H\left(t'-1\right). \tag{74}$$

Proof. First, $\mathbf{D}H(t'+1) < \mathbf{D}G(t'+1)$, implies:

$$\delta\frac{t'}{m}\left[\mathbf{D}V\left(t'+1\right) - \mathbf{D}V\left(t'\right)\right] + \delta\left(1 - \frac{t'+1}{m}\right)\left[\mathbf{D}V\left(t'+2\right) - \mathbf{D}V\left(t'+1\right)\right] < \frac{1}{m}\left[(a-b)-(c-d)\right],$$

It is known that $\mathbf{D}V(t'+1) - \mathbf{D}V(t') > \mathbf{D}V(t'+2) - \mathbf{D}V(t'+1)$. It follows that $\mathbf{D}V(t'+1) - \mathbf{D}V(t') > \mathbf{D}H(t'+1)$. So if it could be shown that $\mathbf{D}V(t'+1) - \mathbf{D}V(t') < \mathbf{D}G(t'+1)$, $\mathbf{D}H(t'+1) < \mathbf{D}G(t'+1)$ would follow as a consequence.

$$\mathbf{D}V\left(t'+1\right) - \mathbf{D}V\left(t'\right) < \frac{1}{m}\left[(a-b)-(c-d)\right],$$

$$\Rightarrow \frac{\frac{1}{m}(a-c) - (1-\delta+\delta/m)\,\mathbf{D}V(t')}{1-\delta\left(1-\frac{t'+1}{m}\right)} < \frac{1}{m}\left[(a-b)-(c-d)\right],$$

Which rearranges to

$$\frac{t'}{m} > \frac{\frac{\delta}{m}(a-c) - \left(1-\delta+\frac{\delta}{m}\right)(c-d) - \frac{\delta}{m^2}\left(1-\delta+\frac{\delta}{m}\right)\left[(a-b)-(c-d)\right]}{\left(1-\delta+2\delta/m+\left(\frac{\delta}{m}\right)^2\right)\left[(a-b)-(c-d)\right]}. \tag{75}$$

It is now obvious that this minimum value for t'/m in (75) is clearly smaller than the minimum value that already exists, in terms of

$$\frac{t'_-}{m} = \frac{\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)\left[(a-b)-(c-d)\right]}, \tag{76}$$

This can be seen from the way in which the numerator of (75) is clearly smaller than that in (76), while the denominator of (75) is clearly larger than the denominator in (76).

So the growth in $H(t)$ must have been less than that of $G(t)$ between t' and $t'+1$. This proves the Lemma.

Second, $\mathbf{D}G(t'-1) > \mathbf{D}H(t'-1)$ implies:

$$\delta\frac{t'-2}{m}\left[\mathbf{D}V\left(t'-1\right) - \mathbf{D}V\left(t'-2\right)\right] + \delta\left(1-\frac{t'-1}{m}\right)\left[\mathbf{D}V\left(t'\right) - \mathbf{D}V\left(t'-1\right)\right] < \frac{1}{m}\left[(a-b)-(c-d)\right], \tag{77}$$

The nature of $\mathbf{D}V(t)$ as an increasing function which increases at an increasing rate means $\mathbf{D}V(t') - \mathbf{D}V(t'-1) > \mathbf{D}V(t'-1) - \mathbf{D}V(t'-2)$. It follows that:

$$\mathbf{D}V\left(t'\right) - \mathbf{D}V\left(t'-1\right) > \mathbf{D}H\left(t'-1\right). \tag{78}$$

So demonstrating that

$$\mathbf{D}V\left(t'\right) - \mathbf{D}V\left(t'-1\right) < \mathbf{D}G\left(t'-1\right), \tag{79}$$

would imply that $\mathbf{D}H(t'-1) < \mathbf{D}G(t'-1)$ and so the Lemma would be proved.

(79) implies:

$$-\frac{\frac{1}{m}(b-d) - (1-\delta+\delta/m)\,\mathbf{D}V(t')}{1-\delta\frac{t'-2}{m}} < \frac{1}{m}\left[(a-b)-(c-d)\right],$$

This can be expressed as a condition on t'/m, as follows:

$$\frac{t'}{m} < \frac{\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)}{\left(1-\delta+2\frac{\delta}{m}+\left(\frac{\delta}{m}\right)^2\right)\left[(a-b)-(c-d)\right]} + \frac{\left[1-\delta+3\frac{\delta}{m}+2\left(\frac{\delta}{m}\right)^2\right]\frac{1}{m}\left[(a-b)-(c-d)\right]}{\left(1-\delta+2\frac{\delta}{m}+\left(\frac{\delta}{m}\right)^2\right)\left[(a-b)-(c-d)\right]}, \tag{80}$$

It is possible now to show that this maximum value for t'/m is greater than the maximum value that already exists from the range discovered earlier, namely $t'_{-}+1$. To see why, consider:

$$\frac{t'_{-}+1}{m} = \frac{\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)}{(1-\delta+2\delta/m)[(a-b)-(c-d)]} + \frac{\frac{1}{m}(1-\delta+2\delta/m)[(a-b)-(c-d)]}{(1-\delta+2\delta/m)[(a-b)-(c-d)]} \tag{81}$$

For notational compactness, let $\theta = (a-b)-(c-d)$. The right hand side of (80) will be greater than the right hand side of (81) if and only if:

$$(1+\delta/m)(1-\delta+2\delta/m)\theta > \delta\left[\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)\right] \tag{82}$$

We now show that, provided $\delta < \delta^{**}$, an expression that is smaller than that on the left hand side of (82) is greater than an expression that is larger than the one on the right hand side of (82). Hence the inequality in (82) must hold where $\delta < \delta^{**}$: Note that:

$$(1-\delta+2\delta/m)\theta < (1+\delta/m)(1-\delta+2\delta/m)\theta \tag{83}$$

$$\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d) > \delta\left[\frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d)\right]. \tag{84}$$

So the inequality in (82) must hold whenever:

$$(1-\delta+2\delta/m)[(a-b)-(c-d)] > \frac{\delta}{m}(a-c) - (1-\delta+\delta/m)(c-d),$$

$$\Rightarrow \delta < \delta^{**} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b) + \frac{1}{m}(b-d)}. \tag{85}$$

So whenever $\delta < \delta^{**}$, the function $H(t)$ will shrink by less than the function $G(t)$ in the move from $t'-1$ to $t'-2$. This ensures that moving from $t'-1$ to $t'-2$, the function $H(t)$ remains above the function $G(t)$. The requirement on δ is of no concern because this is needed anyway to ensure that $t' < m$. ■

Lemma 7, Corollary 8 and Lemma 9 effectively establish a single crossing property of the functions $G(t)$ and $H(t)$. Lemma 6 establishes that the point of the single crossing must be at t', where t' is the sole integer in the range $\left[t'_{-}, t'_{-}+1\right]$. These three lemmas together are sufficient to prove Proposition 5, which establishes the optimal strategy for player 1 which leads to some reputation building.

Proposition 5 *If $\delta \in [\delta^*, \delta^{**}]$, and $a - b > c - d$, then the optimal strategy for player 1 involves building a reputation and playing T when $t < t'$ and consuming out of their reputation and playing B where $t \geq t'$, where t' is the sole integer in the interval $\left[t'_-, t'_- + 1\right]$, and*

$$t'_- = \frac{\delta \left(a - c\right) - m \left(1 - \delta + \delta/m\right) \left(c - d\right)}{\left(1 - \delta + 2\delta/m\right) \left[\left(a - b\right) - \left(c - d\right)\right]} \tag{16}$$

Proof. Follows from Lemma 7, Corollary 8 and Lemma 9. ■

C Additive Separability

Additive separability in the payoffs of player 1 is the special case not yet investigated where $a-b=c-d$. It follows that $a-c=b-d$. The consequences of these two equations can be summarized as follows:

- $a-b=c-d \Rightarrow$ The cost of building a reputation does not depend on the extent to which one is trusted by the short lived players.
- $a-c=b-d \Rightarrow$ The benefit of being trusted is the same irrespective of what player 1 plans to do.

The optimal strategy for player 1 is captured in Proposition 6.

Proposition 6 *When player 1's payoffs satisfy additive separability, player 1 will optimally play T whenever $\delta \geq \delta_s$ and optimally play B whenever $\delta < \delta_s$, where*

$$\delta_s = \frac{a-b}{(a-b)+\frac{1}{m}(b-c)}. \tag{86}$$

Proof. The function $G(t)$ is a weighted average of $a-b$ and $c-d$. When $a-b=c-d$, this weighted average will always have the same value irrespective of the weights (which are what vary with t). So $G(t)=a-b=c-d$ and does not vary with t.

If player 1 always plays B, Lemma 1 applies and the $H(t)$ function does not vary with t, and is given by

$$H(t) = \frac{\frac{\delta}{m}(a-c)}{1-\delta+\delta/m}.$$

So there would be no profitable deviation from always playing B provided that:

$$\frac{\frac{\delta}{m}(a-c)}{1-\delta+\delta/m} < a-b=c-d. \tag{87}$$

If player 1 always plays T, then Lemma 3 applies and the $H(t)$ function does not vary with t and is given by:

$$H(t) = \frac{\frac{\delta}{m}(b-d)}{1-\delta+\delta/m}.$$

There will be no profitable deviation from always playing T provided

$$\frac{\frac{\delta}{m}(b-d)}{1-\delta+\delta/m} \geq a-b=c-d. \tag{88}$$

Since $a - c = b - d$, the left hand side of (88) is the same as the left hand side of (87), and the right hand sides of each condition are clearly identical. However the direction of the inequality is different. So when one condition fails, the other must hold. So playing T in all states is the optimal strategy provided that $\delta \geq \delta_s$, where δ_s can be defined by either (87) or (88) as:

$$\delta_s = \frac{a - b}{(a - b) + \frac{1}{m}(b - c)}. \tag{86}$$

■

The key observation here is that neither the benefit of building a reputation and playing T, nor the benefit of spending what reputation one has and playing B depend on the current state of one's reputation. So the optimal action in one state must be the same as the optimal action in any other state. Hence player 1 will either always play T or always play B.

D Proof of Proposition 8

Proposition 8 *Where $\delta \in [\delta', \delta'']$, and $c-d > a-b$, the optimal strategy is to play $T\ \forall t \geq t''$ and to play $B\ \forall t < t''$, where t'' is the sole integer in the interval $\left(t''_-, t''_- + 1\right)$, where*

$$t''_- = \frac{m\left(1-\delta\right)\left(c-d\right) - \delta\left(b-c\right)}{\left(1-\delta\right)\left[\left(c-d\right)-\left(a-b\right)\right]}. \tag{26}$$

Proof. Define two new functions, $\underline{V}(t)$ and $\overline{V}(t)$ and let these respectively give the continuation value from state t to playing B in all states between state t and state 0; and the continuation value from state t to playing T in all states between state t and state m. By Lemmas 1 and 3, which would apply in each case respectively, these are relatively easy to calculate:

$$\underline{V}(t) = \frac{c}{1-\delta} + \frac{\frac{t}{m}\left(a-c\right)}{1-\delta+\delta/m}, \tag{89}$$

$$\overline{V}(t) = \frac{b}{1-\delta} - \frac{\left(1-\frac{t}{m}\right)\left(b-d\right)}{1-\delta+\delta/m}. \tag{90}$$

The state t'' can be defined by two inequalities: $\overline{V}(t'') > \underline{V}(t'')$ and $\overline{V}(t''-1) < \underline{V}(t''-1)$. The first of these gives:

$$\frac{b}{1-\delta} - \frac{\left(1-\frac{t''}{m}\right)\left(b-d\right)}{1-\delta+\delta/m} > \frac{c}{1-\delta} + \frac{\frac{t''}{m}\left(a-c\right)}{1-\delta+\delta/m},$$

$$\Rightarrow \frac{t''}{m} > \frac{t''_-}{m} = \frac{\left(1-\delta\right)\left(c-d\right) - \frac{\delta}{m}\left(b-c\right)}{\left(1-\delta\right)\left[\left(c-d\right)-\left(a-b\right)\right]}. \tag{91}$$

The second inequality gives $\left(t''-1\right)/m$ as being less than the same fraction. the two inequalities together then prove that t'' must lie somewhere in the interval $\left(t''_-, t''_- + 1\right)$. The condition that $\delta \in (\delta', \delta'')$ ensures that $t''_- \in [0, m]$.

Now to prove that there are no profitable deviations. To prove this, we need to show that $H(t) < G(t)$ holds $\forall t < t''$, and $H(t) > G(t)$ holds $\forall t \geq t''$. The value of $\mathbf{D}V(t)$ is known and is constant $\forall t > t''$ through Lemma 3 and $\forall t < t''$, through Lemma 1. This leads to a value for $H(t)$ that is known and invariant to $t\ \forall t < t''-1$ and $\forall t > t''$.

So we can show that B is indeed the optimal action $\forall t < t''-1$ by showing that $H(t) < G(t)\ \forall t < t''-1$. In this region, we know that

$H(t) = \frac{\delta}{m}(a-c)/(1-\delta+\delta/m)$. We also know that since $G(t)$ is a decreasing function, the smallest value of $G(t)$ will occur at $t = t'' - 2$. Hence we require:

$$\frac{\frac{\delta}{m}(a-c)}{1-\delta+\delta/m} < \frac{t''-2}{m}(a-b) + \left(1 - \frac{t''-2}{m}\right)(c-d),$$

Expressing this as a condition on t''/m, it becomes:

$$\frac{t''}{m} < \frac{(1-\delta)(c-d) - \frac{\delta}{m}(b-c) + \frac{1}{m}(1-\delta+\delta/m)[(c-d)-(a-b)]}{(1-\delta+\delta/m)[(c-d)-(a-b)]} + \frac{\frac{1}{m}(1+\delta/m)[(c-d)-(a-b)]}{(1-\delta+\delta/m)[(c-d)-(a-b)]}. \quad (92)$$

The condition in (92) implies a maximum value for t''/m. To check that the implied maximum value for t'' is indeed greater than the maximum value for t'' that already exists requires the right hand side of (92) to be greater than $(t''_- + 1)/m$:

$$\frac{(1-\delta)(c-d) - \frac{\delta}{m}(b-c) + \frac{1}{m}(1-\delta+\delta/m)[(c-d)-(a-b)]}{(1-\delta+\delta/m)[(c-d)-(a-b)]} + \frac{\frac{1}{m}(1+\delta/m)[(c-d)-(a-b)]}{(1-\delta+\delta/m)[(c-d)-(a-b)]} > \frac{(1-\delta)(c-d) - \frac{\delta}{m}(b-c) + \frac{1}{m}(1-\delta)[(c-d)-(a-b)]}{(1-\delta)[(c-d)-(a-b)]}.$$

Which implies:

$$(1+\delta/m)(1-\delta)[(c-d)-(a-b)] > \delta\left[(1-\delta)(c-d) - \frac{\delta}{m}(b-c)\right], \quad (93)$$

With regards to (93), notice that for the left and right hand sides respectively, note that:

$$(1+\delta/m)(1-\delta)[(c-d)-(a-b)] > (1-\delta)[(c-d)-(a-b)], \quad (94)$$

$$\delta\left[(1-\delta)(c-d) - \frac{\delta}{m}(b-c)\right] < (1-\delta)(c-d) - \frac{\delta}{m}(b-c). \quad (95)$$

Hence, if it can be shown that:

$$(1-\delta)[(c-d)-(a-b)] > (1-\delta)(c-d) - \frac{\delta}{m}(b-c), \quad (96)$$

then (93) would follow. Now (96) would hold if and only if:

$$\delta\left(a-b\right)+\frac{\delta}{m}\left(b-c\right)>a-b,$$
$$\delta>\delta'=\frac{a-b}{\left(1-b\right)+\frac{1}{m}\left(b-c\right)}. \tag{97}$$

Which we know must be the case.

Next we want to show that T is the optimal strategy $\forall t\geq t''+1$. For all such values of t,

$$H\left(t\right)=\frac{\frac{\delta}{m}\left(b-d\right)}{1-\delta+\delta/m}.$$

In order for T to be the optimal strategy, it must be the case that $H\left(t\right)>G\left(t\right)$. The largest value of $G\left(t\right)$, given that it is a decreasing function, will be where $t=t''+1$. This would require:

$$\frac{\frac{\delta}{m}\left(b-d\right)}{1-\delta+\delta/m}>\frac{t''+1}{m}\left(a-b\right)+\left(1-\frac{t''+1}{m}\right)\left(c-d\right),$$

Which gives the condition:

$$\frac{t''}{m}>\frac{\left(1-\delta\right)\left(c-d\right)-\frac{\delta}{m}\left(b-c\right)-\frac{1}{m}\left(1-\delta+\delta/m\right)\left[\left(c-d\right)-\left(a-b\right)\right]}{\left(1-\delta+\delta/m\right)\left[\left(c-d\right)-\left(a-b\right)\right]}. \tag{98}$$

In comparison to the minimum value for t'' established in (91), the numerator of the expression in (98) is clearly smaller and the denominator in (98) is clearly larger. So the minimum value in (98) is smaller than that already established in (91). This is sufficient to prove the Proposition. ■

E Proof of Proposition 10

Proposition 10 *Where the social memory is long enough relative to the payoffs,*

$$m > \frac{(b-d)(c-d)}{(a-b)(a-c)}, \tag{31}$$

the optimal strategy of the long lived player depends on the level of their discount factor, δ as follows:

- $\forall \delta < \delta^{\dagger}$*, play B $\forall t$;*
- $\forall \delta > \delta^{\dagger\dagger}$*, play T for $t = 0, 1$ and play B $\forall t \geq 2$; and*
- $\forall \delta \in \left[\delta^{\dagger}, \delta^{\dagger\dagger}\right]$*, play T for $t = 0$ and play B for $t \geq 1$.*

Where:

$$\delta^{\dagger} = \frac{c-d}{(a-d) - \frac{1}{m}(c-d)}, \tag{32}$$

$$\delta^{\dagger\dagger} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b) + \frac{1}{m}(b-d)}. \tag{33}$$

Proof. The proof proceeds in three stages, proving each of the bullet points in turn. In each case, possible deviations at $t = 0$ and $t = 1$ are the only ones examined. Proposition 9 guarantees that B is the optimal strategy in all states where $t \geq 2$, so there is no need to check for profitable deviations in those states. In each of the two remaining states, what we need to check is the relative levels of $H(t)$ and $G(t)$.

While it is still the case

$$H(t) = \delta \frac{t}{m} \mathbf{D} V(t) + \delta \left(1 - \frac{t}{m}\right) \mathbf{D} V(t+1),$$

the activities of the PR specialist mean that $G(t)$ has changed, so that:

$$G(t) = \begin{cases} c-d & t = 0 \\ a-b & t \geq 1 \end{cases}. \tag{99}$$

E.1 No Reputation

Under this candidate strategy, the value functions in states $t = 0$, $t = 1$, and $t = 2$ will be as follows:

$$V(0) = \frac{c}{1-\delta}, \tag{100}$$

$$V(1) = a + \delta V(1) - \frac{\delta}{m}\mathbf{D}V(1), \tag{101}$$

$$V(2) = a + \delta V(2) - \frac{2\delta}{m}\mathbf{D}V(2). \tag{102}$$

Hence the relevant differences in the value functions are:

$$\mathbf{D}V(1) = \frac{a-c}{1-\delta+\delta/m}, \tag{103}$$

$$\mathbf{D}V(2) = \frac{\frac{\delta}{m}(a-c)}{(1-\delta+\delta/m)(1-\delta+2\delta/m)}. \tag{104}$$

And so:

$$H(0) = \frac{\delta(a-c)}{1-\delta+\delta/m}, \tag{105}$$

$$H(1) = \frac{\frac{\delta}{m}(1+\delta/m)(a-c)}{(1-\delta+\delta/m)(1-\delta+2\delta/m)}. \tag{106}$$

And to ensure there is no profitable deviation at $t = 0$ and $t = 1$ respectively requires:

$$\frac{\delta(a-c)}{1-\delta+\delta/m} < c-d, \tag{107}$$

$$\frac{\frac{\delta}{m}(1+\delta/m)(a-c)}{(1-\delta+\delta/m)(1-\delta+2\delta/m)} < a-b. \tag{108}$$

The first of these can be expressed as:

$$\delta < \delta^{\dagger} = \frac{c-d}{(a-d)-\frac{1}{m}(c-d)}. \tag{32}$$

The left hand side of inequality in (108) is clearly increasing in δ. The numerator is increasing in δ and the denominator is decreasing in δ, so the whole expression must be increasing in δ. To express it in terms of an inequality on δ would result in a quadratic expression, fortunately, deriving the expression directly is, unnecessary. Both (107) and (108) must be satisfied in order for the proposed strategy to be optimal. So if (108) is satisfied at $\delta = \delta^{\dagger}$, then it will also be satisfied $\forall \delta < \delta^{\dagger}$, and so $\delta < \delta^{\dagger}$ will be the requirement to ensure the candidate strategy is indeed optimal.

$$\left. \frac{\frac{\delta}{m}\left(1+\delta/m\right)\left(a-c\right)}{\left(1-\delta+\delta/m\right)\left(1-\delta+2\delta/m\right)} \right|_{\delta=\delta^{\dagger}} < a-b \Leftrightarrow m > \frac{\left(c-d\right)\left(b-d\right)}{\left(a-b\right)\left(a-c\right)}. \tag{109}$$

This proves the optimality of the part of the strategy outlined in the first bullet point.

E.2 Maximum Reputation Building

Under the candidate strategy of playing T in states $t = 0, 1$ and playing B in all other states, the value functions at $t = 0$, $t = 1$, and $t = 2$ will be as follows:

$$V\left(0\right) = d + \delta V\left(1\right), \tag{110}$$

$$V\left(1\right) = b + \delta V\left(2\right) - \frac{\delta}{m}\mathbf{D}V\left(2\right), \tag{111}$$

$$V\left(2\right) = a + \delta V\left(2\right) - 2\frac{\delta}{m}\mathbf{D}V\left(2\right). \tag{112}$$

Hence, taking the relevant differences:

$$\mathbf{D}V\left(2\right) = \frac{a-b}{1+\delta/m}, \tag{113}$$

$$\mathbf{D}V\left(1\right) = \frac{\left(1+\delta/m\right)\left(b-d\right)+\delta\left(1-\frac{1}{m}\right)\left(a-b\right)}{1+\delta/m}. \tag{114}$$

Which means that

$$H\left(0\right) = \frac{\delta\left(1+\delta/m\right)\left(b-d\right)+\delta^{2}\left(1-\frac{1}{m}\right)\left(a-b\right)}{1+\delta/m}, \tag{115}$$

$$H\left(1\right) = \frac{\delta}{m}\left(b-d\right)+\delta\left(1-\frac{1}{m}\right)\left(a-b\right). \tag{116}$$

Ensuring that there are no profitable deviations at $t = 0$ or $t = 1$, then requires that:

$$\frac{\delta\left(1+\delta/m\right)\left(b-d\right)+\delta^{2}\left(1-\frac{1}{m}\right)\left(a-b\right)}{1+\delta/m} > c-d, \tag{117}$$

$$\frac{\delta}{m}\left(b-d\right)+\delta\left(1-\frac{1}{m}\right)\left(a-b\right) > a-b. \tag{118}$$

The second of these conditions rearranges to:

$$\delta > \delta^{\dagger\dagger} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)}. \tag{119}$$

Once again, the condition in (117) could be rearranged to a quadratic, but note that the left hand side is increasing in δ. Setting:

$$A = \frac{\delta\left(1+\delta/m\right)(b-d)+\delta^2\left(1-\frac{1}{m}\right)(a-b)}{1+\delta/m},$$
$$\frac{dA}{d\delta} = \frac{\left(1+\delta/m\right)^2(b-d)+\left(2+\delta/m\right)\left(1-\frac{1}{m}\right)(a-b)}{\left(1+\delta/m\right)^2} > 0,$$

So if the inequality in (117) is satisfied at $\delta = \delta^{\dagger\dagger}$, it will be satisfied $\forall \delta > \delta^{\dagger\dagger}$. Once again, it can be shown that:

$$\left.\frac{\delta\left(1+\delta/m\right)(b-d)+\delta^2\left(1-\frac{1}{m}\right)(a-b)}{1+\delta/m}\right|_{\delta=\delta^{\dagger\dagger}} > c-d \Leftrightarrow m > \frac{(c-d)(b-d)}{(a-b)(a-c)}. \tag{120}$$

This proves the second bullet point of the candidate strategy profile is optimal.

E.3 Partial Reputation

To prove the final part of the strategy, now consider the value functions at $t = 0, 1, 2$, when player 1 plays T at $t = 0$ and plays B in all other states. In that case:

$$V(0) = d + \delta V(1), \tag{121}$$
$$V(1) = a + \delta V(1) - \frac{\delta}{m}\mathbf{D}V(1), \tag{122}$$
$$V(2) = a + \delta V(2) - 2\frac{\delta}{m}\mathbf{D}V(2). \tag{123}$$

Hence the relevant differences in the value function are:

$$\mathbf{D}V(1) = \frac{a-d}{1+\delta/m}, \tag{124}$$
$$\mathbf{D}V(2) = \frac{\frac{\delta}{m}(a-d)}{\left(1-\delta+2\delta/m\right)\left(1+\delta/m\right)}. \tag{125}$$

Thus we can say:

$$H(0) = \frac{\delta(a-d)}{1+\delta/m}, \tag{126}$$

$$H(1) = \frac{\frac{\delta}{m}(a-d)}{1-\delta+2\delta/m}. \tag{127}$$

So ensuring that there are no profitable deviations requires that:

$$\frac{\delta(a-d)}{1+\delta/m} > c-d,$$

$$\Rightarrow \delta > \delta^{\dagger} = \frac{c-d}{(a-d)-\frac{1}{m}(c-d)}. \tag{128}$$

$$\frac{\frac{\delta}{m}(a-d)}{1-\delta+2\delta/m} < a-b,$$

$$\Rightarrow \delta < \delta^{\dagger\dagger} = \frac{a-b}{\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)}. \tag{129}$$

Which is sufficient to prove the Proposition. ■

F The Case Where m is Small

We now consider the possibility that

$$m < \frac{(c-d)(b-d)}{(a-b)(a-c)}.$$

The optimal strategy is recorded in Proposition 11

Proposition 11 *Where the social memory is short relative to the payoffs,*

$$m < \frac{(b-d)(c-d)}{(a-b)(a-c)}, \tag{130}$$

the optimal strategy of player 1 depends on the level of their discount factor as follows:

- $\forall \delta < \underline{\delta}$, *play* B $\forall t$;
- $\forall \delta > \overline{\delta}$, *play* T *for* $t = 0, 1$ *and play* B $\forall t \geq 2$; *and*
- $\forall \delta \in \left[\underline{\delta}, \overline{\delta}\right]$, *play* T *for* $t = 1$ *and play* B $\forall t \neq 1$.

The proof proceeds in a similar way to those that have gone before, though some of the required calculations are much more complex than in the previous proofs. However, before beginning the proof, it is worth highlighting that the relationship between this case and the case where

$$m > (c-d)(b-d) / [(a-b)(a-c)]$$

is similar to the relationship between an increasing and decreasing temptation to cheat on one's reputation when that reputation is unmanaged.

Proof. The proof proceeds in three stages, each associated with one of the bullet points defining the optimal strategy.

F.1 No Reputation

Under these circumstances, the calculation of the continuation values from each state and the differences between them, and so the values of $H(t)$ and $G(t)$ at $t = 0, 1$ proceed just as they did when m was large relative to the payoffs. These result in exactly the same inequalities required to ensure that the candidate strategy profile is an equilibrium strategy profile as when m was large, specifically:

$$\frac{\delta(a-c)}{1-\delta+\delta/m} < c-d, \tag{107}$$

$$\frac{\frac{\delta}{m}(1+\delta/m)(a-c)}{(1-\delta+\delta/m)(1-\delta+2\delta/m)} < a-b. \tag{108}$$

However, now, because m is small relative to the payoffs, it is not the case that (108) necessarily holds whenever (107) does. In fact, the opposite is true, and (108) will hold whenever:

$$\begin{aligned}&\left[(1-1/m)^2(a-b)-\frac{1}{m}(a-b)-\frac{1}{m^2}(b-c)\right]\delta^2\\ &-\left[2(1-1/m)(a-b)+\frac{1}{m}(b-c)\right]\delta+(a-b)>0.\end{aligned} \tag{131}$$

The critical values for the quadratic inequality in (131) can be found via the quadratic formula. For simplicity, let:[13]

$$e=(1-1/m)^2(a-b)-\frac{1}{m}(a-b)-\frac{1}{m^2}(b-c), \tag{132}$$

$$f=-\left[2(1-1/m)(a-b)+\frac{1}{m}(b-c)\right], \tag{133}$$

$$g=a-b. \tag{134}$$

The complication is that we are unable to gauge the sign of the coefficient on δ^2. The first possibility is that this coefficient is positive. In this situation, it is possible to show that $f^2-4eg>0$, so that the roots of the quadratic exist. In which case, the quadratic inequality is satisfied whenever *either* of the following inequalities are satisfied:

$$\delta<\frac{-f-\sqrt{f^2-4eg}}{2e}, \tag{135}$$

$$\delta>\frac{-f+\sqrt{f^2-4eg}}{2e}. \tag{136}$$

Note that $\sqrt{f^2-4eg}<-f$, so both terms will be positive. It can be shown that where $e>0$, $-f/2e>1$, so the inequality in (136) can never be satisfied

[13]The standard parameters for the quadratic formula having already been taken by the payoffs.

by a discount factor. The only remaining remaining way for the quadratic inequality to be satisfied is the through (135) being satisfied. This becomes:

$$\delta < \underline{\delta} = \frac{2\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-c)-\sqrt{\lambda}}{2\left[\left(1-\frac{1}{m}\right)^2(a-b)-\frac{1}{m}(a-b)-\frac{1}{m^2}(b-d)\right]}, \tag{137}$$

where:

$$\lambda = \left[2(a-b)+\frac{1}{m}(b-c)\right]^2 - 4\left(1-\frac{1}{m}\right)(a-b)^2. \tag{138}$$

The remaining possibility is where $e < 0$. In this case, the quadratic inequality is satisfied if and only if *both* of the following inequalities are satisfied:

$$\delta > \frac{-f+\sqrt{f^2-4eg}}{2e}, \tag{139}$$

$$\delta < \frac{-f-\sqrt{f^2-4eg}}{2e}. \tag{140}$$

The inequality in (139) will always be satisfied by any discount factor, since the terms on the right hand side are all negative. So the only remaining inequality to worry about is (140), which is exactly the same as (135). This proves the first part of the optimal strategy.

F.2 Maximum Reputation Building

Once again, the calculations of continuation values and the differences between them $t = 0, 1, 2$ remain exactly as they were in the case where m was large relative to the payoffs. So the functions $H(t)$ and $G(t)$ take the same values as they did in $t = 0, 1$ as in the case where m was large relative to the payoffs. So the maximum reputation building candidate strategy profile will be optimal provided that

$$\frac{\delta(1+\delta/m)(b-d)+\delta^2\left(1-\frac{1}{m}\right)(a-b)}{1+\delta/m} > c-d, \tag{117}$$

$$\frac{\delta}{m}(b-d)+\delta\left(1-\frac{1}{m}\right)(a-b) > a-b. \tag{118}$$

However, because m is small relative to the payoffs, it is not the case that (117) holds whenever (118) does. In fact, the opposite is the case. (117) can be expressed as a quadratic inequality, as follows:

$$\left[\left(1-\frac{1}{m}\right)(a-b)+\frac{1}{m}(b-d)\right]\delta^2+\left[(b-d)-\frac{1}{m}(c-d)\right]\delta-(c-d) > 0. \tag{141}$$

The ensuing discussion is easier if we adopt the following labeling of these parameters:

$$h = \left(1 - \frac{1}{m}\right)(a-b) + \frac{1}{m}(b-d) > 0, \tag{142}$$

$$i = (b-d) - \frac{1}{m}(c-d) < 0, \tag{143}$$

$$j = -(c-d) < 0. \tag{144}$$

The fact that $j < 0$ and $h > 0$ ensures that the roots of the quadratic exist. The quadratic inequality will be satisfied if *either* of the following inequalities is satisfied:

$$\delta < \frac{-i - \sqrt{i^2 - 4hj}}{2h}, \tag{145}$$

$$\delta > \frac{-i + \sqrt{i^2 - 4hj}}{2h}. \tag{146}$$

Since $i > 0$ and $h > 0$, (145) could never be satisfied by a discount factor, since it requires δ to be less than a negative number, which is clearly impossible. So the focus should be on (146), which requires that δ be greater than a particular number. Furthermore, we know that number must be positive, because $h > 0$, and $j < 0$ together imply that $i^2 - 4hj > i^2$, which is enough to ensure that the right hand side of (146) is positive. This becomes:

$$\delta > \overline{\delta} = \frac{\sqrt{\mu} - \left[(b-d) - \frac{1}{m}(c-d)\right]}{2\left[\left(1 - \frac{1}{m}\right)(a-b) + \frac{1}{m}(b-d)\right]}. \tag{147}$$

Where

$$\mu = \left[(b-d) + \frac{1}{m}(c-d)\right]^2 + 4\left(1 - \frac{1}{m}\right)(a-b)(c-d). \tag{148}$$

This completes the proof of the optimality of the second part of the candidate equilibrium strategy for player 1.

F.3 Partial Reputation Building

The candidate strategy profile in the region where $\delta \in \left[\underline{\delta}, \overline{\delta}\right]$ is to play T in state $t = 1$ and to play B in *all* other states, including $t = 0$. In that case, the continuation values in the relevant states are:

$$V(0) = \frac{c}{1-\delta}. \tag{149}$$

$$V(1) = b + \delta V(2) - \frac{\delta}{m}\mathbf{D}V(2). \tag{150}$$

$$V(2) = a + \delta V(2) - 2\frac{\delta}{m}\mathbf{D}V(2). \tag{151}$$

Then the relevant differences in the value functions are as follows:

$$\mathbf{D}V(2) = \frac{a-b}{1+\delta/m}, \tag{152}$$

$$\mathbf{D}V(1) = \frac{(1+\delta/m)(b-c) + \delta\left(1-\frac{1}{m}\right)(a-b)}{(1-\delta)(1+\delta/m)}. \tag{153}$$

It follows that

$$H(0) = \frac{\delta(1+\delta/m)(b-c) + \delta^2\left(1-\frac{1}{m}\right)(a-b)}{(1-\delta)(1+\delta/m)}, \tag{154}$$

$$H(1) = \frac{\frac{\delta}{m}(1+\delta/m)(b-c) + \delta(1-\delta+\delta/m)\left(1-\frac{1}{m}\right)(a-b)}{(1-\delta)(1+\delta/m)}. \tag{155}$$

Ensuring that there is no profitable deviation from these strategies requires that $H(0) \leq c-d$ and that $H(1) \geq a-b$.

The first of these inequalities gives:

$$\frac{\delta(1+\delta/m)(b-c) + \delta^2\left(1-\frac{1}{m}\right)(a-b)}{(1-\delta)(1+\delta/m)} \leq c-d,$$

and a little algebraic manipulation takes this to:

$$\frac{\delta(1+\delta/m)(b-d) + \delta^2\left(1-\frac{1}{m}\right)(a-b)}{1+\delta/m} \leq c-d, \tag{156}$$

The requirement that $H(1) \geq a-b$ becomes:

$$\frac{\frac{\delta}{m}(1+\delta/m)(b-c) + \delta(1-\delta+\delta/m)\left(1-\frac{1}{m}\right)(a-b)}{(1-\delta)(1+\delta/m)} \geq a-b,$$

a little algebraic manipulation will then take this to:

$$\frac{\frac{\delta}{m}(1+\delta/m)(a-c)}{(1-\delta+\delta/m)(1-\delta+2\delta/m)} \geq a-b. \tag{157}$$

At this point, note that (156) is the exact converse of (117) and that (157) is the exact converse of (108). Since (117) could be satisfied by a discount factor if and only if $\delta > \overline{\delta}$, the implication of (156) must be $\delta \leq \overline{\delta}$. Similarly, (108) could be satisfied by a discount factor if and only if $\delta < \underline{\delta}$ so the implication of (157) must be $\delta \geq \underline{\delta}$. It follows that $\delta \in \left[\underline{\delta}, \overline{\delta}\right]$ is sufficient to ensure that the candidate strategy profile is optimal for player 1.

All that remains to be proved is that such a region does indeed exist and that $\underline{\delta} < \overline{\delta}$. One method to check that this is the case would be to directly compare the two expressions. This would be somewhat clunky and is left to the interested reader as an exercise.

A somewhat more elegant method starts from the observation that where

$$m = \frac{(c-d)(b-d)}{(a-b)(a-c)},$$

the inequality from which $\underline{\delta}$ can be derived, (157) held with equality at $\delta = \delta^{\dagger}$, (See previous section). It follows that, when m takes this particular value, $\underline{\delta} = \delta^{\dagger}$. The previous section also demonstrated that when m takes this particular value, the expression from which $\overline{\delta}$ can be derived holds with equality at $\delta = \delta^{\dagger\dagger}$. It follows that, when m takes this particular value $\overline{\delta} = \delta^{\dagger\dagger}$. Finally, it follows from the expressions for $\delta^{\dagger}$ and $\delta^{\dagger\dagger}$ in the previous section that, at this value of m, $\delta^{\dagger} = \delta^{\dagger\dagger}$. Hence, by transitivity, at this value of m, $\underline{\delta} = \overline{\delta}$.

Now note that $\underline{\delta}$ and $\overline{\delta}$ are effectively defined by the following equations:

$$P\left(\underline{\delta}, m\right) = \frac{\frac{\underline{\delta}}{m}\left(1+\underline{\delta}/m\right)(a-c)}{\left(1-\underline{\delta}+\underline{\delta}/m\right)\left(1-\underline{\delta}+2\underline{\delta}/m\right)} = a-b, \tag{158}$$

$$Q\left(\overline{\delta}, m\right) = \overline{\delta}(b-d) + \frac{\overline{\delta}^{2}\left(1-\frac{1}{m}\right)(a-b)}{1+\overline{\delta}/m} = c-d. \tag{159}$$

Standard comparative static arguments would then imply that:

$$\frac{d\underline{\delta}}{dm} = -\frac{\partial P/\partial m}{\partial P/\partial \underline{\delta}}, \tag{160}$$

$$\frac{d\overline{\delta}}{dm} = -\frac{\partial Q/\partial m}{\partial Q/\partial \overline{\delta}}, \tag{161}$$

Focusing first on the maximum value for δ, $\overline{\delta}$, it can be shown that

$$\frac{\partial Q}{\partial m} = \frac{\frac{\overline{\delta}^2}{m^2}\left(1+\overline{\delta}\right)(a-b)}{\left(1+\overline{\delta}/m\right)^2} > 0, \tag{162}$$

$$\frac{\partial Q}{\partial \overline{\delta}} = (b-d) + \frac{\overline{\delta}\left(2+\overline{\delta}/m\right)\left(1-\frac{1}{m}\right)(a-b)}{\left(1+\overline{\delta}/m\right)^2} > 0. \tag{163}$$

It immediately follows that $d\overline{\delta}/dm < 0$ so as m falls, $\overline{\delta}$ increases.

Next, turning to $\underline{\delta}$, it is simple to show that:

$$\frac{\partial P}{\partial \underline{\delta}} > 0, \tag{164}$$

and that

$$\frac{\partial P}{\partial m} < 0. \tag{165}$$

Although this last point requires the explicit assumption that $m \geq 3$ and the use of numerical methods. However, it then follows that $d\underline{\delta}/dm > 0$, and so as m decreases, $\underline{\delta}$ will decrease too. Since

$$\underline{\delta}\Big|_{m=\frac{(c-d)(b-d)}{(a-b)(a-c)}} = \overline{\delta}\Big|_{m=\frac{(c-d)(b-d)}{(a-b)(a-c)}}.$$

and $d\overline{\delta}/dm < 0$ and $d\underline{\delta}/dm > 0$, it follows that:

$$\underline{\delta} < \overline{\delta} \Leftrightarrow m < \frac{(c-d)(b-d)}{(a-b)(a-c)}, \tag{166}$$

Which is sufficient for the purposes of this proposition and so completes the proof. ■