Reputation with Multiple Opponents, and Commitment Bias in Voting

Sambuddha Ghosh

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Economics

Adviser: Dilip Abreu, Roland Benabou, Eric Maskin, Stephen Morris

June 2008

Abstract

The first chapter of my thesis is the first attempt in the game theory literature to study the impact of reputation in a game with an arbitrary number of non- myopic players, not just two. I consider a reputation game with perfect monitoring and *multiple long-lived* opponents indexed $1, 2, \cdots, n$. Player 0, who attempts to build reputation, could be either the normal type (maximizing the discounted sum of payoffs) or one of many commitment types. For a single opponent ($n = 1$) the previous literature finds that, when all players are arbitrarily patient and 0 is patient relative to others, in *any* equilibrium player 0 gets arbitrarily close to g_0^{**}, the maximum feasible payoff for 0 subject to giving player 1 at least her minmax. In other words, 0 can appropriate 'everything'. For $n > 1$, in stark contrast, I find that player 0's minimum equilibrium payoff is typically strictly below g_0^{**} and could be as low as his minmax value.

My first chapter makes two important simplifying assumptions: (i) the type space that 0 can mimic contains only pure strategy types, and (ii) an announcement stage is available before the dynamic game is played. In the second chapter I show that the announcement stage is without loss of generality; although different strategies are called for, the same bound works when player 0 does not have an opportunity to signal his type. This chapter also shows that the qualitative result proved in the first chapter is robust to a very large set of types I extend the result to include types that, in addition to playing arbitrarily complex history-dependant actions, may be committed to mixing. I state and prove a new upper bound on vmin 0 that works even in the presence of types that mix.

My third chapter, co-authored with Vinayak Tripathi, draws attention to a paradoxical commitment bias that may be present in voting games. Consider an electorate whose individual rankings of alternative policies may change between the time they vote for a candidate and the date a policy is implemented. Rankings may change following common or idiosyncratic shocks. Voters choose, via a simple majority election, between a candidate who is committed to a single alternative (no matter what the preferences of the voters are), and an unbiased candidate who implements the ex-post optimum. We show that, even when idiosyncratic shocks lead to slight changes in rankings, the unique symmetric pure strategy Nash equilibrium often entails strategic voters choosing the committed candidate. The resulting welfare loss is higher when a common shock is more likely. The bias is amplified when the voters are strategic.

Acknowledgements

This is perhaps the hardest part, for I cannot do justice in so short a passage to all the influences and people who have enriched my life in ways that I had hardly thought possible. Let me start with my parents and maternal grandparents; I lived with the latter for many years while my parents worked. To all of them I owe my moral and cultural values, and a debt that I cannot repay. I consider myself fortunate to have had great teachers throughout; they were inspiring. The intellectual environment at Princeton, and the people who built and sustained the legacy have all had a role, even if it was indirect. It is an honour to have been a student where Fields medallists and Nobel laureates may be seen ambling around.

My thesis owes a lot to those named below, and any shortcomings that survived did so in spite of them. I am grateful to Dilip Abreu for detailed comments and advice; to Stephen Morris, without whose encouragement this project might not have taken off; to Faruk Gul, Eric Maskin, and Wolfgang Pesendorfer for their suggestions and guidance. Discussions with Satoru Takahashi and Takuo Sugaya helped clarify the proofs; Vinayak Tripathi sharpened my understanding of the literature. Eric Maskins influence extended beyond these papers; I have learnt much from him about doing research, to the extent that this skill may be learnt. The last chapter of my thesis will always remain special because it was my first scholarly paper. It arose from a discussion I had at a cafe with my close friend and classmate Vinayak Tripathi, with whom I had the good fortune to collaborate. Roland Benabou added clarity to our amorphous idea and later practically forced us to present our work, something for which we have not thanked him enough. B.V. Rao patiently taught me most of the Mathematics I know; I will always aspire to the expositional ability of B.V. Rao and Elias Stein. Without Asok Barman of Scottish Church College I would probably have been in a different field. Research in theoretical economics has, perhaps ironically, taught me about life to persevere and to believe in oneself when the chips are down.

An academics life is not quite as dramatic as *The Beautiful Mind* would have us believe. The frisson of a proof that suddenly comes together is rather rare. Ultimately I shall remember Princeton not for the seminars and the scholars, but for the wonderful friendships I forged, cutting across narrow ethnic boundaries. The years I spent at 11 Dickinson Street with various friends will also be among the happiest years of my life, notwithstanding having to wake up for morning classes with curtains of snow coming down. It is here that I met Arudra, Eden (Eddie to most of us), Jenny, John, and Hind. Several other students at Fisher Hall, like Arpita, Konstantinos, Kunal, Mustafa, and Nandita, became good friends through shared experiences. Vinayak has not only been a co-author and a constructive critic, but also someone to rely on. I cannot end this without mentioning two very special friends Jenny and Natalia. The formers advice spanned weight-loss to personal problems; the latters willingness to devour my cooking has been exceeded by her willingness to stand by me when I hit an emotional nadir. Every time I ask myself if I chose the right graduate school, they will always tilt the scales in favour of Princeton.

To my parents, maternal grandparents, and to Asok Barman.

Contents

[1] Co-authored with Vinayak Tripathi.

Chapter 1

Multiple Opponents and the Limits of Reputation: Pure Strategy Types

1.1 INTRODUCTION AND LITERATURE SURVEY

The literature on "reputation" starts with the premise that while we may be *almost* certain of the payoffs and the structure of the game we cannot be *absolutely* certain. The small room for doubt, when exploited by a very patient player (sometimes referred to as a long-run player), leads to "reputation" being built. This paper is the first to investigate what happens when a very patient player who has private information about his type plays a dynamic game with *multiple* non-myopic opponents. The results are significantly different from previous work that has looked at reputation building against a *single* opponent, whether myopic or patient. Even when the long-run player can mimic any type from a very large set and is relatively more patient, my results demonstrate that the presence of multiple opponents imposes strong checks on his ability to exploit reputation.

Kreps and Wilson (1982), and Milgrom and Roberts (1982) introduced this idea in the context of chain-store games with a finite horizon. They show that arbitrarily small amounts of incomplete information suffice to give an incumbent monopolist a rational motive to fight early entrants in a finitely repeated entry game even if in each stage it is better to acquiesce once an entry has occured; in contrast the unique subgame-perfect equilibrium of the complete information game involves entry followed by acquiescence in every period (this is Selten's well-known *chain-store paradox*). Another paper by the above-named four explains cooperation in the finitely repeated prisoner's dilemma using slight incomplete information, although there is no cooperation in the unique subgame perfect equilibrium of the corresponding complete-information game. The first general result on reputation, due to Fudenberg and Levine (1989), henceforth referred to as FL, applies to a long-lived player playing an infinitely repeated simultaneous-move stage-game against a myopic opponent (having a discount factor of 0): As long as there is a positive probability of a type that always plays the Stackelberg action,[1] a suffciently patient LR can approximate his Stackelberg payoff.

Work on reputation has not looked at games with multiple long-lived opponents who interact with one another, not just with the long-run player; my paper investigates how far the results for a single opponent extend to settings with *multiple non-myopic opponents*$(i = 1, 2, \cdots, n; n > 1)$ and one long-run player(0), who could be very patient relative to the rest of the players and has private information about his type. I show that introducing more opponent leads to qualitative differences, not just quantitative ones.

The central issue in the reputation literature is characterising a lower bound on the payoff of player 0 when he has access to an appropriate set of types to mimic

[1] If LR were given the opportunity to commit to an action, the one he would choose is called the Stackelberg action. In other words, the Stackelberg action maximises his utility if his opponent plays a best response to whichever action he chooses.

and is patient relative to the other players. Fudenberg and Levine have shown that this lower bound is "very high" — When a patient player 0 faces a single myopic opponent repeatedly, there is a discontinuity in the (limiting) set of payoffs that can be supported as we go from the complete information to the incomplete information game if we allow a rich enough space of types. Even small ex-ante uncertainties are magnified in the limit, and the effect on equilibria is drastic. Does the same message hold in variants of the basic framework?

Schmidt (1994), Aoyagi (1996), and Celentani et. al. (1994) all consider only one *non-myopic* opponent. At the cost of some additional notation let us be precise about the discount factors $\delta_0, \delta_1, \cdots, \delta_n$. All three papers deal with the limiting case where all players are patient but 0 is relatively patient, i.e. $\delta_i \rightarrow 1\ \forall i > 0$ and $\frac{1-\delta_0}{1-\delta_i} \rightarrow 0\ \forall i > 0$. One standard justification of a higher discount factor for LR is that he is a large player who plays many copies of the same two-player game, while the other party plays relatively infrequently. My work will retain this assumption, although we shall see later that it is not critical to my work and merely facilitates comparison with the earlier work. To see the importance of having a non-myopic opponent, recall FL's argument for a single *myopic* opponent: If the normal type mimics the type that plays the "Stackelberg action" in every period, then eventually the opponent will play a best-response to the Stackelberg action if she is myopic. Schmidt was the first to consider the case of one non-myopic opponent; he shows that this natural modification introduces a twist in the tale: The opponent need not play a period-by-period best response because she fears that she is facing a perverse type that plays like the commitment type on the path but punishes severely off-path. A related problem is that the normal type itself might be induced to behave like the committment type on the path of play, but differently off the path — when 1 deviates, he reveals himself as the normal type and play enters a phase that is bad for 1 but good for the normal type of 0. Since off-path strategies are not learnt in any equilibrium with perfect monitoring, this could lead to very weak lower bounds, ones lower than FL's bound in particular. Schmidt shows that the result in FL extends only to "conflicting interest games" — games where the reputation builder would like to commit to an action that minmaxes the other player. Roughly this is because in such games the impatient player has no choice about how to respond — she must ultimately play her best response in the stage-game or get less than her minmax value, which is impossible in equilibrium. Later Cripps, Schmidt, and Thomas considers arbitary, not just ones with conflicting interest, stage-games and obtains a tight bound that is strictly below the bound of FL; the bound is "tight" in the sense that any payoff slightly higher than the bound can be supported as an equilibrium payoff of player 0.

The subsequent literature argues that Schmidt and Cripps et. al. take a somewhat bleaker view of reputation effects than warranted. In Aoyagi and Celentani et. al. there is no issue whether strategies can be learnt: Trembles in the first paper and imperfect monitoring in the second ensure that all possible triggers are pressed and latent fears do not remain. When strategies are eventually learnt, reputation is once again a potent force. Here is their main result: As the (relatively impatient) opponent also becomes patient in absolute terms, the payoff of the patient player tends towards $g_0^{**} = max\,\{v_0\,|(v_0, v_1) \in F^*\}$,[2] where F^* is the feasible and individually rational set of payoffs; i.e. he gets the most that is consistent with individual rationality of the opponent, player 1.

A final decisive step towards restoring a high lower bound was taken by Evans and Thomas (1997, henceforth ET); they showed that the weakness of the bound in Cripps at. al. relies on the assumption that all irrational types impose punishments of bounded length; under limiting patience they extend the FL result to the case of a single long-lived opponent playing an arbitrary simultaneous stage-game even

[2] We adopt the convention that the generic value is v; w is the "worst" value; and b denotes the best. The mnemonic advantages hopefully justify any break with practice.

under perfect monitoring, provided the reputation builder is patient relative to the opponent. ET obtains the same bound g_0^{**} as Aoyagi and Celentani et. al., showing in the process that a suitable choice of the type space makes the result hold even when there is perfect monitoring. Assign a positive prior probability to a type that plays an appropriate finite sequence of actions, and expects a particular sequence of replies in return; the kth deviation from the desired response is punished punishes for k periods by this type. By mimicking this type, a normal (sufficiently patient) player 0 can approximate his *best feasible payoff* subject to the opponent getting at least her minmax payoff. In terms of the framework used in the literature, my work adds more opponents to the framework used by Schmidt and ET; the results, however, stand in sharp contrast. To focus on a different issue, I shall introduce a signalling stage that abstracts from learning problems that could arise from perfect monitoring. In a later section I show that this modification does not distort results, although it does simplify greatly the description of the strategies.

The immediate generalisation of the bound obtained by Aoyagi, CFLP, and ET for $n = 1$ to multiple ($n > 1$) opponents is $g_0^{**} = max\{v_0 \,|(v_0, v_1, \cdots, v_n) \in F^*\}$. One obvious case where this holds is the one where the relatively patient player (0) is playing a series of independent games with the other n players.[3] When players $1, ...n$ are myopic, the above bound follows readily from the analysis of FL; if the other players are also patient the bound derives from (an n-fold repetition of) the analysis of ET. However I find that this is not true in general: The presence of additional opponents, each with the ability to punish and play out various repeated game strategies, complicates the situation. Recall that the previous literature obtains lower bounds by showing that $v_0^{min} \geq L$, where L is a large lower bound that equals the

[3]By "independent" I mean that the payoff of any player $i > 0$ is independent of the actions of player $j > 0, j \neq i$. This game is no more than the concatenation of n two-player games, each of which has player 0 as one of the two players. Furthermore, the types of player 0 are independent across these games, in the sense that observing the play in any one game conveys no information about the type of 0 in any other game.

Stackelberg payoff of 0 in FL and is g_0^{**} in the three other papers mentioned above. When a player is patient relative to the single opponent he faces repeatedly, reputation is a very powerful force. In contrast, under some non-emptiness restrictions that rule out cases like the immediate extension above, a world with multiple opponents gives a cap on what "reputation" can guarantee 0. In other words, I define a quantity l such that $v_0^{min} \leq l < g_0^{**}$; therefore, l is an *upper bound* on $0's$ minimum equilibrium payoff. I then show that any equilibrium payoff above l of player 0 in the complete information game can be supported in the limiting set of payoffs *even under incomplete information*, and even with any or the types used by Celentani and Schmidt or ET; furthermore, sequential rationality is satisfied in the construction of the above equilibria. Finally, l could be much lower than g_0^{**}, and even as low as the minmax value of player 0. This means that while reputation has non negative value to a player, its impact is qualitatively less dramatic when there are multiple opponents — all equilibrium payoffs of 0 exceeding l in the repeated game of complete information are present in the perturbed game.

Fudenberg and Kreps (1987) is, to my knowledge, the only earlier paper that has *multiple* opponents playing 0, who is trying to build a reputation. Their framework is significantly different from mine; in particular, the impatient players do not have a direct effect one another's payoffs through actions. The paper is concerned with the effect of 0's actions being observed publicly rather than privately by each opponent; their basic "contest" or stage-game is an entry deterrence game, while my analysis considers more general games. At this point I should also note the literature on reputation games in which all players have *equal* discount factors. Cripps et. al. and Chan (2000) have shown that, except in special cases, the reputation builder needs to be patient *relative* to the others to be able to derive advantage from the incomplete information. My main result applies equally well when all players have equal discount factors, although it contrasts most sharply with the previous literature when player

0 is more patient than the rest, which is also the case that has received the most attention in the literature. There is yet another strand of the reputation literature — see for example Benabou and Laroque(1992), and Mailath and Samuelson (2004)— that looks not at the *payoff* from building reputation, but the actual belief of the opponent about the *type* of the reputation-builder. The latter shows that, in a world with imperfect monitoring and a single opponent, even when a patient player can get high payoffs in equilibrium by mimicking a committed type, the opponent will eventually believe with very high probability that she is indeed facing a normal type; in other words, reputation disappears in the long run under imperfect monitoring although the patient player reaps the benefits of reputation. Benabou and Laroque had shown this phenomenon in the context of a particular example. This question will not be relevant in our context.

The plan of the paper is as follows. Section 2 has two motivating examples; section 3 lays out the formal details of the model; the next states the benchmark result for a single patient opponent. Section 5 is the main section, where I prove the upper bound on the minimum equilibium payoffs of 0. I generalise and extend my main result to two additional situations in section 6, although at the cost of increasing complexity of the equilibrium strategies and the analysis. Section 7 chracterises a lower bound on the payoff of player 0; section 8 concludes.

1.2 MOTIVATING EXAMPLES

This section presents a few simple examples to demonstrate the conclusions of this paper. Each example starts with a benchmark case, and in turn illustrates the implications of the theory of repeated games, then of the existing literature on reputation, and finally that of the current paper. They try to weave stories, admittedly simplistic ones, to give a sense of the results. The first example is meant to be the simpler of the two; the second is best postponed until the reader has acquired familiarity with the concepts and definitions introduced in the main model.

Example 1 (A Skeletal Oligopoly Example): A caveat: This example has special features that do not appear in the proof but permit simpler analysis.

Consider a market of size 1 for a perishable good served by three firms $0, 1, 2$; throughout we refer to firm 0 as "he" and use "she" for the others. The demand curve is linear in the total output[4] q:

$$p(q) : \ = \begin{cases} 1 - q \ ; \ q \in [0, 1] \\ 0 \ ; \ q > 1 \end{cases} .$$

There is one *small* side-market of size ϵ shared by firms 1 and 2 : $p(q') := \epsilon - q' \ ; \ q' \in [0, \epsilon]$, where ϵ is a very small positive number and q' is the total output in the side-market. Each market is a Cournot oligopoly at each time $t = 1, 2, \cdots$. For simplicity, there are no fixed and marginal costs. For the infinitely repeated game, firm i discounts profits at rate δ_i; we shall consider the case when all firms are patient but 0 is more so, i.e. all δ_i's are close to 1 but δ_0 is relatively closer. For reasons mentioned in the introduction we add an announcement stage to this game at $t = 0$, allowing player 0 to send a message $m \in \Omega$ about his purported type.

Also note that in the side market of size ϵ the monopoly profit is $\frac{\epsilon^2}{4}$, which is obtained when the total output is $\frac{\epsilon}{2}$; in the Cournot Nash equilibrium (CNE) each firm $i = 1, 2$ produces $\frac{\epsilon}{3}$ and earns a profit of $\frac{\epsilon^2}{9} < \frac{\epsilon^2}{8}$, one-half the monopoly profit.

CASE A: Single Opponent

When there is a single opponent, i.e. firm 2 is absent, the main market is a Cournot duopoly with complete information. In the main market the monopoly profit $\pi^m =$

[4] A firm's output is chosen from a compact convex set; although, strictly speaking, my results deal with finite action spaces, an appropriately fine grid of actions can approximate arbitrarily closely the results that follow.

1/4 is obtained by choosing the output $q^m = 1/2$. The Nash-threats folk theorem of Friedman (1971) shows that the point $(\frac{1}{8}, \frac{1}{8})$, the fair split of the monopoly profit, can be sustained in a subgame-perfect Nash equilibrium (SPNE) if the players 0 and 1 are patient because in the CNE each firm has a strictly lower profit of $\frac{1}{9}$. The following trigger strategy accomplishes it— each firm produces $\frac{1}{4}$, half the monopoly output; any deviation leads to firms switching to the CNE forever.

Now introduce one-sided incomplete information about 0 through a type space Ω, which comprises the normal type ω° and commitment types $\omega_1, \cdots, \omega_K$. The normal type discounts payoffs as usual. However each commitment type $\omega \neq \omega^\circ$ is defined by an output Φ^1_ω for $t = 1$ and for each $t \geq 1$ a mapping Φ^{t+1}_ω from the history of actions of the other players to an output.

We start with the environment of Fudenberg and Levine: $\delta_1 = 0$ (myopic opponent). If there is a positive probability of a type that selects the Stackelberg output irrespective of history, 0 can guarantee himself very close to $\frac{1}{8}$ if he is patient enough. Let us now make even player 1 patient, while 0 is patient relative to her. Evans and Thomas (1997) shows that in now player 0 can get arbitrarily close to his monopoly payoff of $\frac{1}{4}$. Thus the combination of patience and reputation is enough for player 0 to extract the entire surplus— he enjoys (almost) monopoly profits from the main market, while the other player gets (almost) nothing in the main market.

CASE B : Multiple Opponents

Now introduce the third player, making the main market an oligopoly and the side-market a duopoly. Even when we have various commitment types (including the Stackelberg type) we shall see that the monopoly profits in the main market can be split equally among $i = 0, 1, 2$. Quite strikingly, a similar construction can give player 0 *any* positive payoff, not just 1/12.

Why the marked change in results? Let us describe the equilibrium strategies that support this equilibrium. First we describe on-path play. If player 0 declares $m = \omega^o$, then each firm $i = 0, 1, 2$ produces $q_i(t) = \frac{1}{6}$ $\forall i \geq 0$ $\forall t \geq 1$ and makes a profit of $\frac{1}{12}$ in the main market; in the side market play starts with each $i > 0$ producing half the monopoly output and making $\frac{\epsilon^2}{8}$ each. Players $i = 1, 2$ play as follows. Let h^t denote a t-period history; let h^t_+ denote the history of the moves of players 1 and 2. If $m = \omega \neq \omega^o$, players 1 and 2 are called upon at time $t+1$ to produce outputs $\sigma_i^{t+1}(\omega, h^t)$; $i = 1, 2$ such that

$$\Phi_\omega^{t+1}\left(h^t_+\right) + \sum_{i>0} \sigma_i^{t+1}\left(\omega, h^t\right) = 1,$$

i.e. all profits are eliminated in the main market; in the side market each firm $i = 1, 2$ produces $\frac{\epsilon}{4}$ and earns a profit of $\frac{\epsilon^2}{8}$ as before.

Now we specify off-path play:

Case A: $m = \omega^o$: Given that firms are patient, any deviation by 0 may be punished by a reversion to the CNE in the main market, in which each firm produces an output of $\frac{1}{4}$ in the main market and makes a profit of $\frac{1}{16}$; deviations by $i > 0$ are punished by reverting to the CNE in both markets. From the analysis of Friedman we already know that it is an SPNE when all players are patient.

Case B: $m = \omega \neq \omega^o$. If 0 deviates from his announced strategy ever, he reveals himself to be the normal type and, after he is minmaxed to wipe out any gains resulting from the deviation, play moves to the symmetric cooperative equilibrium sustained by Cournot-Nash reversion as in the complete information game. This is an equilibrium if $\delta_i \geq \delta_{ci}$ for some $\delta_{ci} < 1$. A deviation by player $i > 0$ in either market at any point τ impacts play in both markets. In the main market player $j \neq i$ produces a total output of $1 - \Phi_\omega^{t+1}\left(h^t_+\right)$ for all subsequent periods $t + 1 = \tau + 1, \tau + 2, \cdots$, while i is asked to produce 0. In the side market play moves, starting from $\tau + 1$, to the point where firm j produces the Stackelberg output of $\frac{1}{2}\epsilon$ while i responds with her follower output of $\frac{1}{4}\epsilon$; play thus moves to one of the points $P_1 \approx \left(\frac{1}{16}\epsilon^2, \frac{1}{8}\epsilon^2\right)$ or $P_2 \approx \left(\frac{1}{8}\epsilon^2, \frac{1}{16}\epsilon^2\right)$ in the payoff space. In particular, if $i \neq j$ deviates from P_i then play remains at P_i.

Finally we verify that player $i > 0$ does not deviate after any history. The maximum gain from a one-shot deviation is $1/4$, but this leads to a drop in profit in the side-market from $\epsilon^2/8$ to $\epsilon^2/16$. Incentive constraints for both cases A and B are satisfied if

$$(1-\delta_i)\frac{1}{4} + \delta_i \frac{\epsilon^2}{16} \leq (1-\delta_i).0 + \delta_i \frac{\epsilon^2}{8} \implies \delta_i \geq \frac{1}{1+\epsilon^2/4} =: \underline{\delta}.$$

Player i does not deviate from P_i because that is strictly worse: If i deviates she will not be playing a best response to j's output in the current period, and play will remain at P_i. Finally, the normal type of player 0 reveals his type truthfully if

$$(1-\delta_0)\frac{1}{4} + \delta_0.0 \leq \frac{1}{12} \implies \delta_0 \geq \frac{2}{3}.$$

Define $\underline{\delta} := max\left\{\frac{1}{1+\epsilon^2/4}, \frac{2}{3}, \delta_{ci}\right\}$. When $\delta_i \geq \underline{\delta}$, $i = 0, 1, 2$, the above is an SPNE. Here lies the key difference between the single and multiple opponent cases: Mimicking a commitment type gives 0 less than the proposed equilibrium payoff of $\frac{1}{12}$; hence in the SPNE he reveals his type truthfully. ■

Example 2 (Oligopoly with Capacity Constraints) This example is best read after the reader has seen the formal model. Consider an oligopoly with capacity constraints. At each point of time $t = 1, 2, ...$ there is a market of size 1 for a perishable good ; to be specific the demand function is

$$p(q): \ = \begin{cases} 1-q \ ; \ q \in [0,1] \\ 0 \ ; \ q > 1 \end{cases} .$$

There are four firms that serve the market— 0,1,2,3 with capacity constraints $k_i^* = .5, .45, .45, .45$ and zero marginal and fixed costs (adopted for simplicity). 0 is the reputation builder. First restrict attention to the reputation game with $n = 1$. Player $1's$ minmax is $\frac{1}{16}$. So in the complete information repeated game it is possible for 0 to get anything consistent with player 1 getting her minmax. In particular 0 can get $\pi^m - \frac{1}{16} = \frac{3}{16}$. As I argued in example 1 above, player 0 can *guarantee* himself very close to $\frac{3}{16}$ in the incomplete information game by mimicking a "crazy" type. The construction is the same as above; to keep this example short I shall not mention the details of the type space. The assumption on discount factors is also unchanged. Let us now compute the maxminmax of each firm $i > 0$, say firm 1 . This is the minmax value of i when 0 plays an action that is most favourable to i and all other try to minmax her. Suppose that firm 0 produces an output of 0, while firms 2 and 3 produce .45 each. The best firm 1 can achieve is obtained by solving:

$$max_q\{1 - (.9 + q)\}q = .1 * q - q^2$$

The first order condition gives $q^* = \frac{1}{20}$, which generates a profit of $\frac{1}{20} * (\frac{1}{10} - \frac{1}{20}) = \frac{1}{400} = W_i$, the maxminmax profit of i. This srtictly exceeds the minmax value $w_i = 0$ of any $i > 0$, when all other firms flood the entire market; in other words, 0's cooperation is needed to minmax any $i > 0$. Now we need to check that the condition N introduced formally later holds; this requires us to check that no matter what 0 does, the others can find an output vector giving each of them more than $W_i = \frac{1}{400}$. When 0's output it low the other firms find it easier to attain any target level of profit. So pick the worst slice of .5. We wish to find a symmetric point in this slice that gives each player $i > 0$ strictly more than the maxminmax. Find the maximum symmetric point in the slice:

$$max_q \ (\frac{1}{2} - 3q) * q \Rightarrow \frac{1}{2} = 6q^* \Rightarrow q^* = \frac{1}{12}.$$

The associated level of profit for each firm $i > 0$ is $(\frac{1}{2} - \frac{1}{4}) * \frac{1}{12} = \frac{1}{48} > \frac{1}{400}$. Therefore the non-emptiness assumption N that I formally introduce later on is satisfied. Now we have to calculate the lowest profit that 0 could get in any slice subject to the others getting above $\frac{1}{400}$. Even without exact and painstaking calculations it can be shown that this profit is very close to $\frac{1}{100}$; the argument follows. Supose the other firms almost flood the market so that the price is $\frac{1}{50}$; the maximum 0 could be producing is 0.5, earning a profit of $\frac{1}{100}$. Each of the three remaining firms can produce at least $\frac{1}{8}$, thereby making a profit of at least $\frac{1}{8} * \frac{1}{50} = \frac{1}{400} = W_i$. Thus all collusive outcomes of the repeated game that give 0 more than some very low number (below $\frac{1}{100}$) can be sustained even in the repuational game. Recall that in contrast 0 can guarantee himself $\frac{3}{16} \gg \frac{1}{100}$ when only one opponent is present, the presence of multiple opponents bringing about approximately a 20-fold drop in the minimum assured profit of player 0.■

1.3 THE MODEL

There are $n + 1$ players— $0, 1, 2, \cdots, n$; throughout we refer to firm 0 as "he" and to the others as "she". Player 0 is relatively more patient and attempts to build a reputation; we refer to players $i > 0$ as the "opponents". Let us first describe the temporal structure of the complete-information repeated game, which is perturbed to obtain the "reputational" or incomplete-information game. At each time $t = 1, 2, ...$

the following *normal-form* stage-game is played:

$$G = \langle N = \{0, 1, \cdots, n\},\ (A_i)_{i=0}^{n}\ ,\ (g_i)_{i=0}^{n} \rangle\,.$$

N is the player set; A_i is the finite set of pure actions a_i available to player i at each t, while $\mathcal{A}_i := \triangle(A_i)$ is her set of mixed actions α_i. A profile a of (pure) actions all players is in $A := \times_{i=0}^{n} A_i$; the pure action profile a_+ of the players $i > 0$ lies in $A_+ := \times_{i>0} A_i$. A profile of mixed actions α is an element of $\mathcal{A}$. A comment on the use of subscripts is in order: Profiles are denoted without a subscript, the i^{th} element of a profile/vector is denoted by the same symbol but with the subscript i, and the subscript $+$ denotes all players $i > 0$ collectively. The von Neumann-Morgenstern payoff function of agent $i \geq 0$ is $g_i : A \to \mathbb{R}$; and the associated vector payoff function is given by

$$g = (g_0, g_1, \cdots, g_n) : A \to \mathbb{R}^{n+1}.$$

For any $E \subset \mathbb{R}^d$ and any $J \subset \{1, 2, ..., d\}$, the projection of E onto the plane formed by coordinates in J is denoted by E_J:

$$E_J := \{e_J \,|\exists e_{-J} s.t.\, (e_J, e_{-J}) \in E\}\,.$$

The convex hull of any subset E of an Euclidean space is coE; it is the smallest convex set containing E. For any player $i \geq 0$ the *minmax value* w_i and the *pure strategy minmax value* w_i^p are defined[5] respectively as

$$w_i := min_{\alpha_{-i}} max_{a_i} g_i(a_i, \alpha_{-i})\ ;\ w_i^p = min_{a_{-i}} max_{a_i} g_i(a_i, a_{-i}).$$

[5] We adopt the convention that the generic value is v; w is the "worst" value. The mnemonic advantage hopefully justifies any break with practice.

Player i gets her minmax w_i in game G when the action profile m^i is played:

$$g_i\left(m_i^i, m_{-i}^i\right) = max_{a_i \in A_i} g_i\left(a_i, m_{-i}^i\right) = w_i.$$

The feasible set of payoffs in G is

$$F := co\left\{g(a) : a \in A\right\},$$

using an PRD (Public Randomization Device)[6] available to agents; ϕ_t is the observed value of the PRD in period t, and ϕ^t denotes the *vector* of all realized values from period 1 to period t. The weakly individually rational set is $F^* := \left\{v \in F :\ v_i \geq w_i\ \forall i \geq 0\right\}$. Monitoring is perfect.

Players $1, 2, ..., n$ all maximize the sum of discounted per-period payoffs; to reduce notation and facilitate comparison with the literature on repeated games I take a common discount factor for the opponents, i.e. $\delta_i = \delta\ \forall i > 0$.[7] Let us now add incomplete information — the patient player(0) could be one of many types $\omega \in \Omega$; the prior on Ω is given by $\mu \in \triangle(\Omega)$. The type-space Ω contains ω°, the normal type of the repeated game, who maximizes the sum of per-period payoffs using his discount factor δ_0. The other types may be represented as expected utility maximizers, although their utility functions may not be representable as sums of discounted stage-game payoffs. Consider, for example, the "strong monopolist" in the chain-store game of Kreps and Wilson; this type has a dominant strategy in the dynamic game to always fight entry.

[6] Sorin (1986) and Fudenberg and Maskin (1991) showed that a PRD is without loss of generality when players are patient. I continue to make this assumption in the interests of expositional clarity.

[7] Lehrer and Pauzner (1999) look at repeated games with differential discount factors, and find that the possibility of temporal trade expands the feasible set beyond the feasible set of the stage-game. This creates no problems for my results because the feasible set of the stage-game continues to remain feasible even if all δ_is are not equal .

Formally each type $\omega \neq \omega^\circ$ is identified with or defined by the following sequence

$$\Phi^1_\omega \in A_0, \Phi^{t+1}_\omega : A^t_+ \to A_0.$$

This sequence $\Phi_\omega := (\Phi^1_\omega, \Phi^2_\omega, \cdots)$ specifies an initial ($t = 1$) action and for each $t > 1$ maps any history of actions[8] played by players $1, 2, \cdots, n$ into an action of player 0. In what follows we fix (G, Ω, μ), where G is the stage-game, $\Omega = \{\omega^\circ, \omega_1, \cdots, \omega_K\}$ is the set of types, and μ is the prior on Ω. It is to be noted that the above model allows a very rich set of types. The point of the literature on reputation is that the normal type ω° might want to mimic a commitment type $\omega \neq \omega^\circ$ in order to secure a higher payoff. It is clear that the type space Ω must include appropriate types to mimic, a feature captured in reputation papers by means of some form of full support assumption. In order to investigate the maximal impact of reputation, I allow player 0 to mimic strategies with infinite memory as in Evans and Thomas,[9] and strategies with bounded recall as in Celentani et. al.

The dynamic game starts with an announcement phase at $t = 0$: Player 0 sends a message $m \in \Omega$ about his type. Then the repeated game with perfect monitoring is played out over periods $t = 1, 2, \cdots$. Abreu and Pearce (2009) also employed an announcement; that it permits considerable analytical simplicity will be seen from the complexity of the strategies needed when we dispense with the announcement phase. The normal type of player 0 is free to announce a type $m \neq \omega^\circ$, his true type. We assume that commitment types announce truthfully.

EQUILIBRIUM: An equilibrium comprises the following elements, defined recursively over the time index:

(i) a messaging strategy $m : \{\omega^\circ\} \to \Omega$ for the normal type of player 0;

[8] Potentially player 0 could condition his play at t on the PRD up to period t , i.e. on ϕ^t , in addition to a^{t-1}_+. This would not change our results or proofs; the notation is more involved though.

[9] With two players, i.e. one opponent, it is enough (see Aoyagi and Celentani et. al.) to consider all types with bounded recall strategies — each commitment type with bounded recall $\tau \in \mathbb{N}$ plays an action that depends on the actions played by the other player in the past τ periods.

(ii) for each $i = 0, 1, \cdots, n$, a map $\sigma_i(1) : \Omega \times \{\phi^1\} \to A_i$, and a sequence of maps

$$\sigma_i^t : \Omega \times A^{t-1} \times \{\phi^t\} \to A_i;\ t = 1, 2, \cdots \quad \text{for each } t > 1;$$

(iii) common beliefs following the history h^{t-1} of actions up to period $t-1$, denoted by $\mu(.\,|m, h^{t-1}, \sigma) \in \triangle(\Omega)$, are obtained by updating using Bayes rule wherever possible, where m and σ are respectively the equilibrium announcement function and strategy profile.

To simplify we shall often use μ_{t-1} for the above beliefs and omit the variable ϕ^t from the domain of the period t strategies. Additionally we stipulate that for $\omega \neq \omega^o$ the strategy σ_ω^t is given by Φ_ω as long as he has not deviated from the latter.[10] This defines a commitment type. Let $\sigma_\omega := \{\sigma_\omega^t\}_{t \geq 1}$. The strategy σ_i of player $i > 0$ is the sequence of maps $\{\sigma_i^t\}_{t \geq 1}$; Σ_i denotes the set of all strategies of player i. As usual, a strategy profile is

$$\sigma := (\sigma_0, \sigma_1, ..., \sigma_n) \in \Sigma := \Sigma_0 \times \Sigma_1 \times \cdots \times \Sigma_n.$$

In what follows $u_i(\ )$ refers to the discounted value of a strategy profile to player i, possibly contingent on a certain history of actions and an announcement; thus $u_i(\sigma\,|m, h^{t-1}, \mu_{t-1})$ is the normalized sum of the future payoffs of player i discounted to period t, following announcement m and the history h^{t-1}, given that players play according to the strategy profile σ and beliefs are μ_{t-1}. We refer to the payoff of the normal type of player 0 whenever we use $u_0(\ )$. We now state the following familiar definition.

[10]Strategies need to be specified for all histories including ones off the path. By the very definition of a commitment type there are no histories where type ω has violated his strategy, hence what actions we specify here are not important. We assume that any type ω maximizes the discounted sum of payoffs after a history that is not consistent with the commitment type ω.

DEFINITION: A tuple $\left(m^*, (\sigma_0^*, \sigma_1^*, \cdots, \sigma_n^*), \{\mu(.\,|m, h^{t-1}, \sigma^*)\}_{t>1}\right)$ defines a Perfect Bayesian Equilibrium (PBE) if no player has a strictly profitable unilateral deviation:

(i) $u_0(\sigma\,|m^*(\omega^\circ)) \geq u_0(\sigma\,|\hat{\omega})\,\forall\hat{\omega} \in \Omega$, i.e. the messaging strategy m^* is optimal for player 0;

(ii) beliefs μ_{t-1} are obtained by updating the prior using Bayes rule on the equilibrium path and off-path are restricted so that beliefs (about the type of player 0) cannot be affected by actions of players $i \neq 0$;

(iii) for any $i \geq 0$ and for any message ω and any $t-1$ period history h^{t-1} we also have

$$u_i\left(\sigma^*\left|m, h^{t-1}, \mu^{t-1}\right.\right) \geq u_i\left(\hat{\sigma_i}, \sigma_{-i}^*\left|m, h^{t-1}, \mu^{t-1}\right.\right)\ \forall\ \hat{\sigma}_i \in \Sigma_i;\ \forall\ m \in \Omega;\ \forall i \geq 0.$$

My main result involves constructing a truthtelling PBE, as will be seen in section 5.

1.4 REPUTATION RESULTS FOR A PATIENT OPPONENT

With the above notation in place I state and summarise the result with a single long-lived opponent and perfect monitoring, due to Evans and Thomas(ET). This result is the counterpart for $n = 1$ of my result, and is therefore the right benchmark, which stands in stark contrast to my main result in the next section. ET makes the following simplifying assumption:

Assumption PAM *(Pure Action Minmax) : 0 can minmax 1 by playing a pure action.*

This assumption, while restrictive, is adopted for technical simplicity; otherwise mixed strategies need to be learnt as in Fudenberg and Levine (1992). Suppose we wish to approximate the best payoff g_0^{**} for 0 to at most a margin ϵ of error. Find a sequence of action profiles/pairs $(a_0^{**}(t),\ a_1^{**}(t))_{t=1,\cdots,T}$ such that 0's average discounted payoff over this block of T action pairs is close to g_0^{**} , while $1's$ average discounted payoff exceeds her minmax value:

$$\left|\frac{1}{T}\sum_{t=1}^{T} g_0\left(a_0^{**}(t),\ a_1^{**}(t)\right) - g_0^{**}\right| < \epsilon/3 \text{ and } \frac{1}{T}\sum_{t=1}^{T} g_1\left(a_0^{**}(t),\ a_1^{**}(t)\right) > w_1.$$

Let $\hat{\omega}$ be the type that plays as follows:

(a) 0 starts by playing the block of T actions $(a_0^{**}(t))_{t=1}^{T}$;

(b) if player 1 responds with $(a_1^{**}(t))_{t=1}^{T}$, he repeates this block;

(c) when player 1 deviates for the k^{th} time from playing the role prescribed above, player 0 minmaxes her for k periods using the pure strategy minmax from PAM above;

(d) 0 returns to step (a) irrespective of what actions player 1 responded with during punishment.

The key feature is that the type $\hat{\omega}$ of 0 metes out harsher punishments if 1 continues to deviate. Define $V(\delta_0, \delta_1) \subset \mathbb{R}^2$ as the set of Bayes-Nash equilibrium payoffs for discount factors δ_0 and δ_1 respectively. The associated payoff set for player 0 only is given by the projection $V_0\left(\delta_0, \delta_1\right) \subset \mathbb{R}$ of this set onto dimension 0. It might be useful to remind the reader of the following notation, which we introduced earlier — g_0^{**} is the maximum feasible payoff of LR consistent with player 1 getting at least her minmax.

Proposition 0 (Evans and Thomas, 1997) : *Suppose PAM holds and* $\mu\left(\hat{\omega}\right) > 0$*, i.e. the prior* μ *places a positive weight on* $\hat{\omega}$*. Given* $\epsilon > 0$ *there exists a* $\delta_1^{min} < 1$ *such that for any* $\delta_1 > \delta_1^{min}$*, we have* $lim_{\delta_0 \to 1}\ inf V_0\left(\delta_0, \delta_1\right) > g_0^{**} - \epsilon$.

Proof Sketch: See ET for details; a sketch follows.

Fix $\epsilon > 0$; this is the margin of error we shall allow in approximating the best payoff g_0^{**}.

Step 1 : We have seen that the block of action profiles/pairs $(a_0^{**}(t),\ a_1^{**}(t))_{t=1,\cdots,T}$ has the property that $\frac{1}{T}\sum_{t=1}^{T} g_0\left(a_0^{**}(t),\ a_1^{**}(t)\right)$, is within $\epsilon/3$ of g_0^{**} , while 1's average payoff $\frac{1}{T}\sum_{t=1}^{T} g_1\left(a_0^{**}(t),\ a_1^{**}(t)\right)$ is greater than her minmax value.

Step 2 : Consider the type $\hat{\omega}$ defined above. If $\mu(\hat{\omega}) > 0$, one available strategy of the normal type of 0 is to declare and mimic $\hat{\omega}$; the payoff from doing this is a lower bound on his equilibrium payoff.

Step3 : Apply Lemma 1 of ET (the *Finite Surprises Property* of FL) to show that if player 0 follows the above strategy at most a certain finite number of punishment phases can be triggered without player 1 believing that with a high probability he is facing the type $\hat{\omega}$. Also note that since punishments get progressively tougher, the on-path play gets almost as bad as being mixmaxed forever, whereas for a patient player 1 the discounted per-period payoff from $(a_0^{**}(t),\ a_1^{**}(t))_{t=1,\cdots,T}$ exceeds her minmax value.

Together these two observations imply that in any Nash equilibrium a patient player 1 must eventually (after triggering enough rounds of punishment) find it worthwhile to experiment with the actions $(a_1^{**}(t))_{t=1,\cdots,T}$. Once she does so, by construction 0 gets a *mean* payoff which is within $\epsilon/3$ of $\frac{1}{T}\sum_{t=1}^{T} g_0\left(a_0^{**}(t),\ a_1^{**}(t)\right)$. If δ_0 is close to 1, then the discounted and undiscounted payoffs are very close:

$$\left|\sum_{t=1}^{T} g_0\left(a_0^{**}(t),\ a_1^{**}(t)\right) - \sum_{t=1}^{T} \delta_0^t . g_0\left(a_0^{**}(t),\ a_1^{**}(t)\right)\right| < \epsilon/3$$

The average *discounted* payoff is therefore within $2\epsilon/3$ of g_0^{**}. Finally we re-use the fact that 0 is very patient — If 0 is relatively patient, losses sustained while mimicking

$\hat{\omega}$ cannot cost him more than another $\epsilon/3$ in terms of payoffs; thus mimicking type $\hat{\omega}$ assures player 0 payoffs within ϵ of g_0^{**}. ■

The upshot is that it is possible to secure very high payoffs for the normal type of 0 in the incomplete information game even when his opponent is also patient, as long as 0 is relatively patient and has the option to mimic types that punish successive deviations with increasing harshness.

1.5 THE MAIN RESULT : UPPER BOUND FOR $n > 1$

This section contains the main result, showing that in a reputational game with multiple long-lived opponents the minimum equilibrium payoff of 0 is bounded away from his best payoff g_0^{**}. Proposition 2 also characterizes the gap between the bound and g_0^{**}. The following concepts are introduced— *slice*, *conditional minmax*, and *maxminmax*. Any choice of an action $a_0 \in A_0$ by player 0 induces a game among the n opponents and the associated $(n+1)$-dimensional payoff set is called *slice*-a_0; actions in A_0 thus have a one-to-one relation with slices.

DEFINITION: The game $G(a_0)$ is induced from G where player 0 is restricted to use action a_0:

$$G(a_0) := \left\langle N = \{1, \cdots, n\}; (A_i)_{i>0}; (\hat{g}_i)_{i>0} \right\rangle$$

such that $\hat{g}_i(a_+) = g_i(a_0, a_+) \, \forall a_+ \in A_+$.

DEFINITION: Slice-a_0 is the *conditionally feasible set*[11] of payoffs in the induced

[11] This terminology is meant to suggest that conditional on player 0 playing a_0, this is the feasible set of payoffs.

game $G(a_0)$:

$$F(a_0) := co\{g(a_0, a_+) : a_+ \in A_+\} \subset \mathbb{R}^{n+1}$$

Notice that the conditionally feasible set lies in the payoff space of $n+1$ players although only n players have non-trivial moves in any slice.

DEFINITION: The *conditional minmax* of $i > 0$ in the slice-a_0 is the minmax of $i > 0$ *conditional* on $G(a_0)$; call it $w_i(a_0)$.

This is defined as the minmax of player i in $G(a_0)$ among players $1, \cdots, n$, rather than in the game G among $n+1$ players. The set of mixed strategy profiles of players $i = 1, \cdots, n$ is denoted by

$$\mathcal{A}_+ := \mathcal{A}_1 \times \cdots \times \mathcal{A}_n, \text{ where } \mathcal{A}_j := \triangle A_j.$$

Now we define the conditionally minmaxing punishment of player $i > 0$ in slice-a_0 as $m^i(a_0) \in \mathcal{A}_+$ such that

$$g_i\left(a_0, m^i_{-i}(a_0), m^i_i(a_0)\right) = max_{a_i}\ g_i\left(a_0, m^i_{-i}(a_0), a_i\right) = w_i(a_0).$$

Denote $g_j(a_0, m^i(a_0)) = w^i_j(a_0)$ as the payoff to player $j \neq i$ when i is being conditionally minmaxed in slice-a_0.

DEFINITION: The *maxminmax* of a player $i > 0$ with respect to 0 is defined as the maximum among all conditional minmaxes

$$W_i := max_{a_0 \in A_0} w_i(a_0).$$

It is like the usual minmax but under the additional restriction that when others ($j \neq i$ and $j > 0$) try to minmax $i > 0$, player 0 takes the action that is best for i. In general the *maxminmax* is strictly greater than the *minmax* value[12] w_i, in as much as the others require the active cooperation of player 0 to punish i most severely. Truncate the set $F(a_0)$ below W_i to get

$$\mathcal{F}(a_0) \ := \{v \in F\,(a_0) : \ v_i \geq W_i \ \forall \ i > 0\} .$$

Both sets $F\,(a_0)$ and $\mathcal{F}\,(a_0)$ are compact subsets of $\mathbb{R}^{n+1}$; the projection of $\mathcal{F}\,(a_0)$ onto the 0^{th} coordinate is denoted by the subscript 0, i.e. $\mathcal{F}_0\,(a_0) \subset \mathbb{R}$. The worst payoff for player-0 in the slice-a_0 subject to each player i getting above her maxminmax is

$$w_0\,(a_0) := min \ \mathcal{F}_0\,(a_0) \equiv min\,\{v_0 : (\ v_0, v_+) \in \mathcal{F}(a_0)\} .$$

Now consider the maximum of these worst payoffs (one for each slice) for player 0:

$$l \ := max_{a_0 \in A_0} \ w_0\,(a_0)\,;$$

if each $w_0(a_0) < w_0$ set $l = w_0$. This is the maximum among the worst payoffs for player 0, one for each slice, subject to all others getting at least their maxminmax. We now introduce a non-emptiness assumption**:**

Assumption N**:** *$W \in \bigcap_{a_0 \in A_0}\{\mathcal{F}_+\,(a_0)\}$,which has a non-empty interior.*

N (for Non-emptiness) implies the existence of a point that is inside each slice and slightly dominates the maxminmax vector W for players $i > 0$. Recall that F^* is the feasible individually rational set of the complete information game. A full-

[12] $w_i := min_{\mathcal{A}_0} w_i(\alpha_0). \equiv min_{\mathcal{A}_{-i}} max_{A_i} u_i(\alpha_{-i}, \ a_i)$

dimensionality assumption, first introduced in FM, facilitates comparison with the complete information folk theorems:

Assumption FD **(*Full Dimensionality*):** *The set* F^* *has dimension* $n+1$.

***LEMMA 1*:** Assumptions N *and* FD *imply that* $g_0^{**} > l$.

***PROOF*:** See the appendix. ■

We now state and prove the threat points lemma (TPL), which is useful in establishing the bound below. This lemma states that in every slice a_0 we can find an action profile $\rho(a_0)$ for players $i > 0$ giving each $i > 0$ strictly greater than her W_i and player 0 "close" to $w_0(a_0)$ in slice a_0.

THREAT POINTS LEMMA **(*TPL*):** *Fix* $\epsilon > 0$. *There exists a map* $\rho : A_0 \to \mathcal{A}_+$ *such that* $\forall a_0 \in A_0$

$$g_i(a_0, \rho(a_0)) > W_i \ \forall i > 0 \ \text{and} \ g_0(a_0, \rho(a_0)) < w_0(a_0) + \epsilon \leq l + \epsilon.$$

PROOF **(*TPL*):** Pick any slice a_0. Since $\mathcal{F}(a_0)$ is compact, it follows from the definition of $w_0(a_0)$ that

$$\exists \ x_1, x_2, \cdots, x_n \text{ s.t. } (w_0(a_0), x_1, x_2, \cdots, x_n) \in \mathcal{F}(a_0).$$

By N we can pick $\bar{x}$ such that $\bar{x} \in \mathcal{F}_+(a_0) \forall a_0$ and $\bar{x}_i > W_i$ for all $i > 0$. Since $\bar{x} \in \mathcal{F}_+(a_0)$, for each a_0 we can find $\bar{x}_0(a_0) \in \mathbb{R}$ such that $(\bar{x}_0(a_0), \bar{x}) \in \mathcal{F}(a_0)$. Now for any $p \in [0,1]$, the convexity of $\mathcal{F}(a_0)$ implies that the convex combination

$$v(p) := (1-p)(w_0(a_0), x) + p(\bar{x}_0(a_0), \bar{x}) \in \mathcal{F}(a_0) \forall a_0.$$

Pick p small enough so that $(1-p).w_0(a_0) + p.\bar{x}_0(a_0) < w_0(a_0) + \epsilon$, which is always possible because $\epsilon > 0$ and the LHS tends to $w_0(a_0)$ as $p \to 0$. Since $v(p) \in \mathcal{F}(a_0)$, we can find $\rho(a_0) \in \Delta(A_+)$ s.t. $g(\rho(a_0)) = v(p)$. This ρ satisfies the statement of the lemma. ■

For any $i \geq 0$, let $\lambda_i(a_0) := g_i(a_0, \rho(a_0))\,; a_0 \in A_0,$ and then define the minimum of these payoffs across all slices.

$$\lambda_i := min_{a_0}\{\lambda_i(a_0)\}.$$

Thus $i > 0$ gets at least λ_i in any slice if all players $j > 0$ play according to $\rho(\,)$ and 0 mimics the announced type. Since $\lambda_i(a_0) > W_i\ \forall a_0$, we also have $\lambda_i > W_i$. Proposition 2 below shows that $v_0^{min}(\delta_0, \delta_1, \cdots, \delta_n)$ can be brought arbitrarily close to l. The proof, which is constructive, assumes that mixing is ex-post observable; this is a technically convenient assumption. The argument hinges on our ability to force the relatively very patient player 0 to reveal himself at the announcement stage, i.e. to use $m(\omega^\circ) = \omega^\circ$ at time 0.

How do we ensure that ω° does not want to deviate and announce a commitment type ω? If the normal type is announced, play the complete information equilibrium $\sigma^{ci}(l + 3\epsilon)$, giving player 0 a payoff of $l + 3\epsilon$; here ϵ is an arbitrarily small positive quantity. If a commitment type $\omega \neq \omega^\circ$ is announced, players $i > 0$ believe that they are indeed facing type ω. If he is indeed type ω, player 0 must play according to $\Phi(\omega)$; after the history h^{t-1}, he must induce slice-$\Phi_\omega^t\left(h_+^{t-1}\right)$ at time t. From TPL above we can construct a "threat point" in each slice according to $\rho(\,)$. Players $i > 0$ use $\Phi(\omega)$ to keep track of player 0's actions, and respond to h^{t-1} with $\rho\left(\Phi_\omega^t\left(h_+^{t-1}\right)\right)$. From TPL we know that this gives player 0 no more than $l + \epsilon$, while each player $i > 0$ gets strictly more than her maxminmax value W_i. If 0 has never deviated and some $i>0$ deviates, the deviating player is punished by conditional minmaxing, i.e. she is minmaxed in the slice that 0 plays. (Since 0 has not deviated he will follow $\Phi(\omega)$.) The proof shows that this deters cheating by any $i>0$.

Given that players $i>0$ don't deviate, 0 cannot profit by announcing $\omega \neq \omega^\circ$ and following $\Phi(\omega)$ because he loses at least 2ϵ relative to declaring his true type. If he announces $\omega \neq \omega^\circ$ and then deviates, he immediately exposes himself as the normal type. This is followed by an equilibrium where all players put probability 1 on the normal type and play the complete information equilibrium where 0 gets $l+2\epsilon$. This still falls ϵ short of announcing truthfully at $t=0$; so a patient 0 will never declare ω and then reveal himself subsequently. He will follow the equilibrium strategies off-path too: if he has deviated and announced $\omega \neq \omega^\circ$ he strictly prefers to reveal himself at the first opportunity. In short, the constructed equilibrium gives 0 less than $l+2\epsilon$ if he is ω° but announces $m \neq \omega^\circ$. So he prefers to announce ω° truthfully and get $l+3\epsilon$.

***PROPOSITION 2*:** *Fix $\epsilon > 0$. When $n > 1$, under assumptions N and FD and (ex-post) observable mixed strategies, we have $v_0^{min}(\delta_0, \delta_1, \cdots, \delta_n) \leq l+\epsilon$ if each δ_i is high enough, where $l < g_0^{**}$.*

***PROOF*:** Fix $\epsilon > 0$ to be small enough that $l+3\epsilon \in F_0^*$, where F_0^* is the projection of F^* onto player 0's payoff space. By TPL we have $\rho : A_0 \to \triangle(A_+)$ such that

$$g_0(a_0, \rho(a_0)) \leq l+\epsilon \ \forall a_0 \in A_0 \text{ and } W_i < \lambda_i := min_{a_0}\{\lambda_i(a_0)\} \, \forall i>0.$$

Recall that for any i and any a_0, we defined $\lambda_i(a_0) := min_{a_0}\{g_i(a_0, \rho(a_0))\}$. By N,

$\exists \eta \in \mathbb{R}^n$ and $\Delta > 0$ such that[13]

$$\eta \in int\left(\mathcal{F}_+\left(a_0\right)\right) \forall\, a_0 \text{ and } W_i < \eta_i < \eta_i + \Delta < \lambda_i.$$

For each $j > 0$ define the vector $\eta(j) \in \mathbb{R}^n$ by

$$\eta_j(j) = \eta_j \text{ and } \eta_i(j) = \eta_i + \Delta < \lambda_i \ \forall i \neq j.$$

Since η lies in the interior of $\mathcal{F}_+\left(a_0\right)$, if Δ is small we have

$$\eta(i) \in \mathcal{F}_+\left(a_0\right) \forall i > 0 \text{ and } \forall a_0 \in A_0.$$

Finally, let $\sigma^{ci}\left(v_0\right)$ denote an equilibrium (SPNE) strategy profile of the complete information repeated game that gives ω° a payoff of v_0, and gives players $i > 0$ any feasible payoffs $v_i > w_i$. We assert that the following strategies are part of a PBE giving ω^0 a payoff of $l + 3\epsilon$.

EQUILIBRIUM STRATEGIES:

On-path play is as follows. Player 0 reveals his type by making the announcement $m(\omega) = \omega$. If the announced type at $t = 0$ is ω°, then by Bayes rule we get $\mu\left(\omega^\circ|\omega^\circ\right) = 1$; from $t = 1$ onwards play the equilibrium strategy profile $\sigma^{ci}(l+3\epsilon)$. If $m = \omega \neq \omega^\circ$, update beliefs to $\mu(\omega|\omega) = 1$ and start *Phase I*. In this phase the prescribed play at t conditional on the $(t-1)$-period history h^{t-1} is $\rho\left(\Phi^t_\omega\left(h^{t-1}_+\right)\right) \in \triangle\left(A_+\right)$, which gives player 0 an expected payoff $w_0\left(\Phi^t_\omega\left(h^{t-1}_+\right)\right) + \epsilon \leq l + \epsilon$ if he plays according to Φ^t_ω.

Off-path play is as follows. Suppose player 0 has never deviated from $\Phi(\omega)$ in the past and that player $i > 0$ deviates at time τ; then play enters *Phase* $II(i)$, where

[13] Pick an η that lies in a small open ball centered at W and in each $\mathcal{F}_+\left(a_0\right)$.

player i is *conditionally minmaxed* in slice $\Phi_{\omega}^{t}\left(h_{+}^{t-1}\right)$ at time $t=\tau+1,...,\tau+P$, where the length P of the punishment satisfies the following inequality:

$$PW_i + max_{a\in A}\ g_i(a) < P\eta_i + min_{a\in A}\ g_i(a)\ \forall i > 0....(*)$$

This condition is always satisfied for some large enough integer P since $W_i < \eta_i$. During this phase, player i is asked to play a conditional (pure) best response to $m_{-i}^{i}\left(\Phi_{\omega}^{t}\left(h_{+}^{t-1}\right)\right)$. Once $Phase\ II(i)$ is over play moves into $Phase\ III(i)$.

In $Phase\ III(i)$ at each $t \geq \tau+P+1$, if the history is h^{t-1} play the action profile $\alpha_{+}(i)$ such that

$$g_{+}\left(\hat{a}_0, \alpha_{+}(i)\right) = \eta(i)\ \text{, where } \hat{a}_0 = \Phi_t(\omega)\left(h_{+}^{t-1}\right).$$

This gives and expected payoff vector $\eta(i)$ in this phase; by construction $\eta(i)$ incorporates a small reward of Δ for each player j other than the last player(i) to deviate. If player j unilaterally deviates from $Phase\ II(i)$ or $Phase\ III(i)$ then impose $Phase\ II(j)$ followed by $Phase\ III(j)$, and so on.

If type ω° has announced $m=\omega\neq\omega^{\circ}$ and mimicked type ω till time $\tau-1$, resulting in history $h^{\tau-1}$, he reveals himself at τ by deviating from Φ_{ω} by picking a best response $a_0(\tau)$ to $\rho\left(\Phi_{\omega}^{\tau}\left(h_{+}^{\tau-1}\right)\right)$ from within the actions that the announced type ω would not take at τ:

$$a_0(\tau) \in argmax\ g_0\left(a_0, \rho\left(\Phi_{\omega}^{\tau}\left(h_{+}^{\tau-1}\right)\right)\right)\ \text{ s.t. }\ a_0 \neq \Phi_{\omega}^{\tau}\left(h_{+}^{\tau-1}\right).$$

If player 0 deviates for the first time at τ resulting in the history h^{τ}, set $\mu(\omega^{\circ}|\omega, h^{\tau}) = 1$ and from time $\tau+1$ start playing $\sigma^{ci}\left(l+2\epsilon\right)$.

VERIFICATION OF EQUILIBRIUM:

After $m = \omega \neq \omega^\circ$ has been announced, we need to check that every player $i > 0$'s strategy is unimprovable holding the strategy of the others to be fixed when 0 has never deviated before (this is enough because if 0 has deviated in the past if we are in a complete information game, where deterring deviations is well-understood.)

Step 1: Player $i > 0$ doesn't deviate from playing according to ρ in *Phase I*. Her minimum payoff on the path is λ_i, while her maximum payoff from a one-shot deviation is $(1-\delta) M + \delta \left(1-\delta^P\right) W_i + \delta^{P+1} \eta_i$, which converges to η_i as δ goes to 1. By construction $\eta_i < \lambda_i$.

Step 2: Player $i > 0$ does not deviate from $Phase\ II(j)$ or $Phase\ III(j)$ because she ends up with η_i rather than $\eta_i + \Delta$.

Step 3: Player $i > 0$ does not deviate from $Phase\ II(i)$ because she is anyway playing her best response in each slice; any other action else gives a lower payoff in the current period and also prolongs the punishment; deviating from $Phase\ III(i)$ is not profitable because inequality $(*)$ ensures that restarting the punishment is costly for i.

Step 4: We now reason that ω° does not deviate if $\delta_0 \geq M/(M+\epsilon)$.

(a) If $(1-\delta_0)M < \delta_0 \epsilon$, he will not find it profitable to announce $\omega \neq \omega^\circ$ and reveal himself later.

(b) Consider the off-path situation where ω° has announced $m \neq \omega^\circ$ at time 0. If he reveals himself by a deviation he gets $l + 2\epsilon$ in the continuation game whereas mimicking m gives him no more than $l + \epsilon$. Since $(1-\delta_0)M < \delta_0 \epsilon$, the maximum cost of revealing himself is outweighed by the gain ϵ in the continuation equilibrium.

Therefore ω° announces truthfully. The result follows since ϵ is arbitrary. ■

***COROLLARY*:** *If* $(v_0, v_+) \in \mathbb{R}^{n+1}$ *is an equilibrium payoff vector under complete information and* $v_0 > l$*, then the payoff* v_0 *is an equilibrium payoff of* ω° *in any reputational game* (Ω, μ) *when all players are patient:*

$$[v_0 \in V_0^{ci} \text{ and } v_0 > l\ ] \Rightarrow v_0 \in V_0.$$

The next proposition concerns the payoffs of *all* players, not just 0. It is a quasi folk-theorem result showing that any equilibrium payoff vector v of the complete information game with limiting patience can be supported as an equilibrium payoff of the perturbed game if (i) the probability of the normal type is high enough, (ii) $v_0 > l$ and (iii) $v_i > w_i\ \forall i > 0$ (rather than requiring v_i to exceed the maxminmax W_i).

***PROPOSITION 3*:** *Assume N, FD, and observable mixing. Fix and strictly individually rational payoff* v *of the complete information game. If* $v_0 > l$*, for any sequence of priors* $\left(\mu^k\right)_{k\in\mathbb{N}}$ *such that* $\mu^k(\omega^\circ) \to 1$ *we can construct a sequence of equilibria* $\left(\sigma^k\right)_{k\in\mathbb{N}}$ *in the games* $\left(\Omega, \mu^k\right)$ *satisfying*

$$u\left(\sigma^k\right) := \left(u_0\left(\sigma^k\right), \cdots,\ u_n\left(\sigma^k\right)\right) \to v.$$

***PROOF*:** Fix $(v_0, v_+) \in V^{ci}$ such that $v_0 > l$. Construct a truth-telling equilibrium σ^k of the reputational game $\left(\Omega, \mu^k\right)$ such that $u_0\left(\sigma^k\right) = v_0$. If the announced type is ω°, then play $\sigma^{ci}(v_0)$; if $m \neq \omega^\circ$ then play the equilibrium outlined in Proposition 2 where the normal type gets less than $(v_0 + l)/2$. (Take ϵ of Proposition 2 as any positive no. below $(v_0 - l)/2$.) Note the inequality

$$\mu^k(\omega^\circ) \cdot v_+ + \left(1 - \mu^k(\omega^\circ)\right) \cdot M \geq u_+\left(\sigma^k\right) \geq \mu^k(\omega^\circ) \cdot v_+ + \left(1 - \mu^k(\omega^\circ)\right) \cdot (-M).$$

Our proposition now follows since the upper and the lower bounds on $u_+\left(\sigma^k\right)$ above converges to same quantity v_+ as $\mu^k(\omega^\circ) \to 1$. ■

REMARK: Note that neither this proposition nor the one before can be obtained as a consequence of the folk theorem for stochastic games proved in Dutta (1996).

1.6 SOME EXTENSIONS

This section extends the main result to two additional situations of interest. The earlier section makes an assumption that is critical for the proof to work — probabilites of mixing by any player are ex-post perfectly observable. While this might be a reasonable assumption in some circumstances, one can equally easily come up with ones where this is less natural. This creates a problem during the punishment phases of players $i > 0$. Note that punishing player 0 for declaring a non-normal type does not make use of the observability of mixed strategies, because conditional on the realisation of a public signal each i plays a pure action. Punishing any player $i > 0$ is problematic when one cannot observe how $j \neq i, 0$ mixed: If player $j \neq i, 0$ is not indifferent between the myopic payoffs of all the actions in the support of m_j^i , then j will not mix with the desired probabilities when minmaxing player i in Phase $II(i)$; deviations to actions outside the support of m_j^i are readily detected and deterred in the usual way. The first subsection below addresses this limitation. The second subsection extends the result to games where the long run player moves first.

1.6.1 UNOBSERVABLE MIXED STRATEGIES

While the pure strategy minmax does not need mixing to be observable, it is a less severe punishment than the usual mixed-action minmax, and can support only a smaller set of payoffs in equilibrium. However FM notes that even when mixing is not observable we can use the minmax as punishment against players $i > 0$ in *Phase II*(i); the trick is to adjust the continuation values at the end of the minmax phase depending on the *realised* sequence of actions of player j during the minmax phase so that $j > 0, j \neq i$ indifferent between all actions in the support of m_j^i.

Proposition 3 : *Fix* $\epsilon > 0$. *Assume N and FD hold. Even when mixing by* $i > 0$ *is* unobservable, *we have* $v_0^{min} \leq l$.

Proof:[14] Fix $\epsilon > 0$. It is enough to show that there exists a payoff of the incomplete information game in which player 0 gets $\hat{v}_0 := l + \epsilon$ and any $i > 0$ gets at least W_i. The following quantities are as in the proof of Proposition 1: $\eta \in \mathbb{R}^n$ and $\Delta > 0$ such that $W < \eta < \eta + \overrightarrow{\Delta} < \lambda$,[15] where $\overrightarrow{\Delta} := (\Delta, \cdots, \Delta)$, $\lambda := (\lambda_1, \cdots, \lambda)$, and $W := (W_1, \cdots, W_n)$; P is defined as before by $(*)$ and $\rho(.)$ is the same as in TPL.

If 0 declares himself as normal type(i.e $m = \omega^\circ$), play the (complete information) repeated games equilibrium $\sigma^{ci}(\hat{v}_0)$ giving $\hat{v}_0$ to 0, and at least W_i to all others. If $m \neq \omega^\circ$, then start play in *Phase I*. In describing the phases below we repeatedly use terms of the form z_j^i. Suppose we order the actions in the support of $m_j^i(a_0)$ in increasing order of the expected utility they give when 0 plays a_0 and the player $k \neq j, k > 0$ play $m_k^i(a_0)$. Let $p_j^i(K)$ denote the amount by which the expected utility of the K^{th} action in the support exceeds that of the first action. Now define

$$z_j^i := \frac{1 - \delta_j}{\delta_j^P} \sum_{s=\tau}^{\overline{\tau}} \delta_j^{s-1} p_j^i(K(s)),$$

[14] To prevent cluttering of notation we do not make explicit the presence of the PRD in our proofs.

[15] We define vector inequalities as follows: $x < y \Leftrightarrow x_i < y_i \forall i$; $x \leq y \Leftrightarrow x_i \leq y_i \forall i$ and $x \neq y$; $x \leqq y \Leftrightarrow x_i \leq y_i \forall i$.

where $K(s)$ is the action that j actually played in the s^{th} period of the relevant phase.[16] After player i has been minmaxed or conditionally minmaxed, we transition to a point adjusted by the quantities z_j^i. Also note that for high δ the magnitude of z_j^i cannot exceed $\Delta/2$; this ensures that no continuation value below goes outside the feasible set.

Phase I: In period t play the action n-tuple $\rho\left(\Phi_t(m)\left(h_+^{t-1}\right)\right)$, where h_+^{t-1} is the history of actions upto and including period $t-1$. If player 0 deviates from his announced strategy during any phase, go to *Phase* $II(0)$. Suppose that 0 has never deviated from his announced type. If player $i > 0$ deviates unilaterally from any phase, switch to *Phase* $II(i)$: Conditionally minmax i for P periods. Then go to *Phase* $III\left(i, z_1^i, z_2^i, \cdots, z_n^i\right)$, where $\left(\eta_1 + \Delta - z_1^i, \cdots, \eta_i, \ldots \cdot, \eta_n + \Delta - z_n^i\right)$ is the expected payoff vector to players $i > 0$ at $t \geq \tau + P + 1$. The i^{th} component of is η_i, and all other components $j > 0$ are $\eta_j + \Delta - z_j^i$; $\eta(i)$ incorporates a small reward of Δ for each impatient player j other than the last player(i) to deviate, and z_j^i is the adjustment that takes away any excess xepected reward and makes player j indifferent across all the strategies in the support of the mixed action that he needs to play to punish $i > 0$. If player j unilaterally deviates from *Phase* $II(i)$ or *Phase* $III(i)$ then impose *Phase* $II(j)$ followed by *Phase* $III(j)$, and so on. Consider the other case— type ω° has announced $m = \omega \neq \omega^\circ$. He is instructed to play according to $\Phi(m)$. Suppose his first deviation is at at time τ when we are in *PhaseI* or $II(i)$ or $III(i)$ (for some $i > 0$); if the resulting in history h^τ, we set $\mu\left(\omega^\circ \left|\omega, h^\tau\right.\right) = 1$; player 0 is then minmaxed for enough periods to wipe out gains, followed by a switch to the complete information equilibrium $\sigma^{ci}\left(\lambda_0\left(\underline{a}_0\right), \lambda_1\left(\underline{a}_0\right) - z_1^0, \cdots, \lambda_n\left(\underline{a}_0\right) - z_1^0\right)$.

From the standard folk theorem argument it follows immediately that the proposed

[16]If 0 has not revealed rationality then the the support of the actions in $m_j^i\left(a_0\right)$ and their ranking in increasing order of expected utility will vary with a_0. Note that in FM the set of actions is the same at all $s = \tau, \cdots, \bar{\tau}$, whereas they could and in general do vary when we are in the incomplete information game above. If 0 has revealed rationality and is also playing a mixed strategy m_0^i instead of a pure action, then the above definition is appropriately adjusted.

strategies constitute a sequentially rational equilibrium following the announcement $m = \omega^\circ$. All that remains is to verify that the play following $m \neq \omega^\circ$ is also sequentially rational, and that the normal type of player 0 has an incentive to tell the truth. This is done in a number of steps, checking that there is no history such that a unilateral one-step deviation by any $j \geq 0$ at the chosen history gives j strictly greater utility in the continuation game than the proposed equilibrium strategies. That this suffices follows from the definition of NE and well known results in dynamic programming. We first check the incentive constraints for the impatient players and then for 0.

Step 1: $i > 0$ has no incentive to deviate from $Phase\ I$.

If i deviates the maximum gain is $(1-\delta)\, b_i + \delta \left(1-\delta^P\right) w_i + \delta^{P+1}\eta_i$, which is less than v_i for δ close to 1 because it converges to η_i in the limit and by construction $\eta_i < v_i$.

Step 2: $i > 0$ has no incentive to deviate from $Phase\ II(j)$, where $j \neq i$.

If i deviates to an action outside the support then i's per-period payoff in the game converges to $\eta_i < \eta_i + \Delta/2$ in the long run. Thus she does not get the reward $\Delta/2$, which is given for carrying out the punishment. Given the definition of z_i^j , player i's utility is independent of the probabilities of mixing.

Step 3: i has no incentive to deviate from $Phase\ II(i)$.

If i deviates to an action outside the support, she not only plays a suboptimal response in the current period but also re-starts the punishment; this lowers the current and future utility stream.

Step 4: $i > 0$ has no incentive to deviate from $PhaseIII(j)$, where $j \neq i$.

Step 5: i has no incentive to deviate from $Phase\ III(i)$.

If i deviates the maximum gain is $(1-\delta)b_i + \delta\left(1-\delta^P\right) w_i + \delta^{P+1}\eta_i$. The payoff from conformity to the equilibrium is η_i. Thus a suffcient condition to rule out any profitable deviations is

$$(1-\delta)b_i + \delta\left(1-\delta^P\right)w_i + \delta^{P+1}\eta_i < \eta_i$$

As $\delta \to 1$ the LHS converges to the RHS. Rearranging, the above is equivalent to

$$b_i + \delta\left(1+\cdots+\delta^P\right)w_i < \left(1+\cdots+\delta^P\right)\eta_i.$$

As $\delta \to 1$ the LHS converges to $b_i + Pw_i < (P+1)\eta_i$, which is what the RHS converges to. This holds from the definition of P .

Step 6: We now reason that ω° does not deviate. If he announced $m \neq \omega^\circ$ at time 0, then he follows $\Phi(\omega)$ because he gets a lower payoff by deviating and then playing the worst possible slice $\underline{a}_0$ always: He gets $\lambda_0\left(\underline{a}_0\right) = min_{a_0}\lambda_0\left(\underline{a}_0\right)$ at each time. The normal type annouces ω° truthfully because his payoff is $\hat{v}_0 = l+\epsilon$ when he announces truthfully and sticks to the equilibrium, whereas it is less than or equal to $\hat{v}_0$ if he either announces anything else and faithfully mimics that type or if he announces $\omega \neq \omega^\circ$ and does not play like the announced type. ■

1.6.2 LONG – RUN PLAYER MOVES FIRST

This section extends the results of the previous section to the case where the long-run player moves first. Given any simultaneous stage game G as above, define the extensive form stage-game G_{seq} in which player 0 moves first and players 1,2,...n move simultanously after observing the action chosen by player 0. There is the obvious and natural one-to-one mapping from the set of action $(n+1)$- tuples of G to the set of terminal nodes of G_{seq} ; use that to define utilites for G_{seq} .

Corollary 4 : *Even without announcements there exist sequentially rational equi-*

libria giving (normal)player 0 a payoff of $l+\epsilon$ when the stage game is of the form G_{seq} and N holds.

Proof: same as in Proposition 1 above. ■

1.7 A LOWER BOUND FOR n > 1

Having already seen that the LR player can be forced down to payoffs arbitrarily close to l , we end by looking at what payoffs LR can actually guarantee himself. Start by considering types that have a dominant strategy to play a constant action a_0 in every period (bounded recall zero). Fix any $a_0 \in A_0$, and ask the question: What can the LR player guarantee himself by mimicking a type that plays this constant action each period: $a_0^t = a_0 \; \forall \; t = 1, 2, ...$? First define the corresponding individually rational slice $F^*(a_0)$ defined as follows, where w_i is the minmax of i in G:

$$F^*(a_0) := \{v \in F(a_0) \; : \; v_i \geq w_i \forall \; i > 0\}.$$

Thus both $\mathcal{F}(a_0)$ and $F^*(a_0)$ are defined by truncating the slice-feasible payoff set $F(a_0)$ below some level for each player $i > 0$. In the first case this level is the maxminmax for each player, and in the second case this is the conditional minmax. Note that $F^*(a_0) \subset \mathbb{R}^{n+1}$, the $n+1$-dimensional Euclidean plane. Take the infimum of the projection of this set onto dimension 0 ; that is the lowest payoff that 0 can get in an equilibrium if he sticks to a_0 for all t. The rough reasoning is that if 0 continues to play a_0 , the others cannot continue to play a strategy profile that gives them less than their respective minmax values. If all he could do was to mimic a type that plays a constant action every period (a bounded recall strategy with 0 memory) the worst he could do is to get the max over $a_0 \in A_0$ of these minima. Thus we have a lower bound $l_0 := max_{a_0 \in A_0} inf \; F_0^*(a_0)$; in any equilibrium player-0 cannot get much less than l_0 when all players are patient and he is relatively patient, and the prior places positive weight on all types that play a constant action a_0 for all t. A formal statement follows. We make the following assumption:

Assumption TC: All "crazy" types declare their type truthfully.

This assumption, it should be remarked, is not assuming truthtelling for the entire game. In no way does it constrain the normal type ω°'s announcement.[17] This assumption has been used by Abreu and Pearce (2002, 2007) as a shortcut to an explicit model with trembles or imperfect monitoring, in which the strategies would eventually be learnt whether or not some irrational types declare truthfully. However such a model would of necessity be technically challenging to handle, especially in view of the large type space I wish to support. One might thus justify it as contributing to technical simplicity.[18]

Next, a richness assumption will capture the premise of reputational arguments— Even when we believe that a player is overwhelmingly likely to be of a given type, we can never be *absolutely* sure that he is. Naturally reputational arguments are interesting precisely because they work even when one type — the normal type of the corresponding repeated game— $\omega^\circ \in \Omega$ has high probability mass, i.e. $\mu(\omega^\circ) \approx 1$. The following makes the assumption that the prior μ places a positive but arbitrarily small weight on all types that play a constant action every period.

Assumption ACT (**A**ll **C**onstant action **T**ypes) **:** $\mu(\omega(a_0)) > 0 \ \forall a_0 \in A_0$.

Under this rather weak assumption ACT and TC we have the following result that puts a lower bound on 0's payoff across all BN equilibria.

[17] If truthtelling is to hold there it must be derived; otherwise studying reputation is pointless.

[18] A section of the next chapter extends the analysis to situations where announcements are unavailable; the accompanying proof would not need TC .

Proposition 5 (Lower Bound): *Under TC and ACT there exists a lower bound* $l_0(\delta_0, \delta_1)$ *on the payoffs on* 0 *in a BN eq. such that* $lim_{\delta \to 1} lim_{\delta_0 \to 1} l_0(\delta_0, \delta) = l_0$ *, where* $l_0 := max_{a_0 \in A_0} inf \; F_0^*(a_0)$.

Proof: This proof is omitted because it is standard in the existing literature; see for example Cripps, Schmidt, and Thomas. ■

The above step establishes the existence lower bound on 0's min payoff in BN eq.[19] using strategies that involve playing a single action at all times. Note that the definition of $F^*(a_0)$ uses the minmax value w_i , which would in general be less than a conditional minmax defined earlier.

In a game where strategies can be learnt because of trembles or imperfect monitoring the set $F^*(a_0)$ would be replaced by the set $\mathcal{H}(a_0)$, which uses the conditional minmax rather than the minmax: $\mathcal{H}(a_0) := co\{v \in F(a_0) : \; , v_i \geq w_i(a_0) \forall \; i > 0\}$. This would raise the lower bound as $w_i(a_0^*) > w_i$ in general. We use $F^*(a_0)$, which truncates $F(a_0)$ below the minmax w_i rather than below the conditional minmax $w_i(a_0)$, because there might exist equilibria in which even if player 0 plays a_0 always it is not clear to the others that this is the case; consequently they perceive their lowest eq. payoff as w_i rather than $w_i(a_0)$.

In general, it is clear that constant action types constitute a small class of strategies; there are uncountably infinite ways of switching between the various actions in A_0, each action amounting to the choice of a "slice" of the game G. The next step would be to look at increasingly longer strategies of bounded recall, perhaps using some kind of induction on the memory size and see to what extent they lead to further improvements. This unfortunately turns out to be a very hard problem to solve, partly because there are uncountably many "crazy" strategies that 0 could

[19] This bound thus applies to all BNE, not just the smaller subset of PBE.

potentially mimic. In principle one could think of each "crazy" strategy as a rule for transitioning among the slices G ; thus solving the incomplete information game is akin to solving a class of stochastic games among n rather than $n+1$ players, defined using the repeated game. We then need to find for each game in the class the worst possible payoff of the normal type, and finally taking sup/max over all these possible minima(or infima). Since stochastic games are notoriously hard to handle, this compounds the difficulty. However as we have seen, LR cannot guarantee himself anything more than l no matter how patient he is relative to his opponents, who are also patient.

1.8 CONCLUSION

This paper contributes to the literature on reputation that starts with the work of Kreps, Wilson, Milgrom and Roberts, and is sharpened into powerful and general theoretical insights by Fudenberg and Levine, and subsequent papers. My paper analyses reputation formation against *multiple patient opponents*. I show that there are some additional insights to be gained from this case, over and above the elegant theoretical insights of the previous literature with a single opponent. While reputation is in general valuable even against multiple players, it may not be possible for the patient player to extract the entire surplus while leaving the others with barely their *minmax* values. Let v_0^{min} be the minimum equilibrium payoff of player 0 in the limit when all players are patient and 0 is patient relative to the rest. I find an *upper bound* l such that $v_0^{min} \leq l$: Any payoff of the (complete information) repeated game in which 0 gets more than l can be sustained. A single opponent cannot threaten credibly to punish and thwart a patient player trying to build reputation. But with more than one patient opponent, there might be ways to commit to punishing even a patient player for not behaving like the normal type.

Bibliography

[1] ABREU, D : “On the Theory of Infinitely Repeated Games with Discounting”, Econometrica, Vol. 56, No. 2 (1988)

[2] ABREU, D , AND P. K. DUTTA, AND L SMITH: “The Folk Theorem for Repeated Games: A Neu Condition”, Econometrica, Vol. 62, No. 4 (1994)

[3] ABREU, D , D PEARCE ; Bargaining, Reputation and Equilibrium Selection in Repeated Games with Contracts; Econometrica

[4] AOYAGI, M ; Reputation and Dynamic Stackelberg Leadership in Infinitely Repeated Games , Journal of Economic Theory, 1996, vol. 71, issue 2, pages 378-393

[5] BENABOU, R; G LAROQUE: Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility; 1992, The Quarterly Journal of Economics, vol. 107(3)

[6] CELENTANI, M ; D FUDENBERG ; D. K. LEVINE ; W PESENDORFER : Maintaining a Reputation Against a Long-Lived Opponent; Econometrica, Vol. 64, No. 3 (May, 1996)

[7] CHAN, J : On the Non-Existence of Reputation Effects in Two-Person Infinitely-Repeated Games, April 2000, Johns Hopkins University working paper.

[8] CRIPPS, M. W. , E. DEKEL, W. PESENDORFER : Reputation with Equal Discounting in Repeated Games with Strictly Conflicting Interests, Journal of Economic Theory, 2005, vol. 121, 259-272.

[9] CRIPPS, M. W. ; G. J. MAILATH; L SAMUELSON [CMS] : Disappearing Reputations in the Long-run — Imperfect Monitoring and Impermanent Reputations Econometrica, Vol. 72, No. 2. (Mar., 2004), pp. 407-432.

[10] CRIPPS, M. W. ; G. J. MAILATH; L SAMUELSON : Disappearing private reputations in long-run relationships ; Journal of Economic Theory, Volume 134, Issue 1, May 2007, Pages 287-316

[11] CRIPPS, M. W. , K M. SCHMIDT and J P. THOMAS : Reputation in Perturbed Repeated Games; Journal of Economic Theory, Volume 69, Issue 2, May 1996, Pages 387-410

[12] CRIPPS, M. W. , and J P. THOMAS : "Some Asymptotic Results in Discounted Repeated Games of One-Sided Incomplete Information," Mathematics of Operations Research, 2003, vol. 28, 433-462.

[13] DUTTA, P. K. : A folk theorem for stochastic games; Journal of Economic Theory, 1995

[14] EVANS, R. ; J. P. THOMAS : Reputation and Experimentation in Repeated Games With Two Long-Run Players Econometrica, Vol. 65, No. 5. (Sep., 1997), pp. 1153-1173.

[15] D FUDENBERG ; D. K. LEVINE : Reputation and Equilibrium Selection in Games with a Patient Player , Econometrica, Vol. 57, No. 4 (Jul., 1989)

[16] D FUDENBERG ; D. K. LEVINE : Maintaining a Reputation when Strategies are Imperfectly Observed; The Review of Economic Studies, Vol. 59, No. 3 (Jul., 1992)

[17] FRIEDMAN J. : A Non-cooperative Equilibrium for Supergames James W. Friedman The Review of Economic Studies, Vol. 38, No. 1 (Jan., 1971), pp. 1-12

[18] D FUDENBERG ; D. M. KREPS : Reputation in the Simultaneous Play of Multiple Opponents ; The Review of Economic Studies, Vol. 54, No. 4 (Oct., 1987)

[19] D FUDENBERG ; E. MASKIN : The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, Econometrica, Vol. 54, No. 3. (May, 1986), pp. 533-554.

[20] KREPS, D ; R WILSON : Reputation and Imperfect Information ; Journal of Economic Theory, 1982

[21] KLAUS M. SCHMIDT : Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests; Econometrica, Vol. 61, No. 2 (Mar., 1993)

Appendix

Proof of TPL: STEP 1: $g_0^{**} \geq l$. For any $a_0 \in A_0$, $\exists a_+$ such that $g_0(a_0, a_+) = w_0(a_0)$ and $g_i(a_0, a_+) \geq W_i \forall i > 0$. Since $W_i \geq w_i$, either $g(a_0, a_+) \in F^*$ or $w_0(a_0) < w_0$. In either case $w_0(a_0) \leq g_0^{**} := \text{Max } \{v_0 \,|(v_0, v_+) \in F^*\}$. Since $l := max\{w_0(a_0)\}$ and a_0 is arbitrary, it follows that $g_0^{**} \geq l$.

STEP 2: $g_0^{**} \neq l$. Notice that

$$g_0^{**} = l \implies \exists a_0' \in A_0 \text{ such that } w_0(a_0') = l := max_{a_0} w_0(a_0).$$

Consider $\mathcal{F}_+(a'_0)$, which has a non empty interior by N. There cannot be a point x in $\mathcal{F}(a'_0)$ where $x_0 > g_0^{**}$ because this would contradict the definition of g_0^{**}. So $v_0 \leq g_0^{**} = l$ for any $v \in \mathcal{F}(a'_0)$. Now we show that the above inequality is actually an equality. Suppose that $\exists v \in \mathcal{F}(a'_0)$ with $v_0 = l$; i.e. $\exists \bar{a}_+ \in A_+$ such that $g_0(a'_0, \bar{a}_+) < l = g_0^{**}$. Since $\mathcal{F}_+(a'_0)$ has a non empty interior we can take a point in its interior and convexify it with $g(a'_0, \bar{a}_+)$ such that the convex combination $\bar{v}$ is in $\mathcal{F}(a'_0)$ and $\bar{v}_0 < l$, leading to the contradiction that $l = w_0(a'_0) \leq \bar{v}_0 < l = g_0^{**}$. So we must have $g_0(a'_0, a_+) = g_0^{**}\ \forall a_+ \in A_+$, which implies $w_0 \geq g_0^{**}$. The reverse inequality follows from the definition of these quantities. Therefore $w_0 = g_0^{**}$, which implies $dim(F^*) \leq n$ and therefore contradicts FD.

Steps 1 and 2 taken together imply that $g_0^{**} > l$. ■

Chapter 2

Reputation with Multiple Opponents: Mixed Strategy Types

2.1 INTRODUCTION

The literature on "reputation" perturbs a complete information game by introducing a small probability that one player could be committed to various dynamic game strategies; this small room for doubt leads to a dramatic change in the equilibrium payoffs when a sufficiently patient player with private information about his type exploits it to build "reputation". Work on reputation has not studied cases where *multiple* non-myopic opponents interact with one another, not just with the relatively patient player; my work is the first[1] to investigate how far the results for a single opponent extend to settings with *multiple non-myopic opponents*$(i = 1, 2, \cdots, n; n > 1)$ and one relatively patient player(0), who attempts to build reputation.

The basic issue in the reputation literature is: How high a payoff can 0 guarantee

[1] As mentioned in the introduction to the earlier chapter of this thesis, Fudenberg and Kreps (1987) is the only paper to my knowledge that has multiple opponents playing 0, who is trying to build a reputation. However their framework is significantly different from mine; in particular, the opponents do not affect one another's payoffs through actions. They are concerned with the effect of 0's actions being observed publicly rather than privately by each opponent when the basic "contest" is an entry deterrence game.

himself when he has access to an appropriate set of types to mimic and is patient relative to the other players? The first general result on reputation, due to Fudenberg and Levine (1989), henceforth FL, applies to a long-lived player 0 playing an infinitely repeated simultaneous stage-game against a myopic opponent: As long as there is a positive probability of a type that plays the Stackelberg action[2] at each time, a sufficiently patient reputation-builder can approximate his Stackelberg payoff. When a patient player 0 faces a single myopic opponent, allowing for a rich enough space of types leads to a discontinuity in the (limiting) set of equilibrium payoffs as we go from the complete information to the incomplete information game: Even small ex-ante uncertainties are magnified in the limit, and the effect on equilibria is drastic. Does the same message hold in variants of the basic framework?

Reputation was introduced into the literature by Kreps and Wilson (1982), and Milgrom and Roberts (1982) in the context of finite-horizon chain-store games. They showed that arbitrarily small amounts of incomplete information suffice to give an incumbent monopolist a rational motive to fight early entrants in a finitely repeated entry game even if in each stage it is better to acquiesce once an entry has occurred; in contrast the unique subgame-perfect equilibrium of the complete information game is marked by entry followed by acquiescence in every period (this is Selten's well-known *chain-store paradox*). Another paper by the above-named four explains cooperation in the finitely repeated prisoner's dilemma using slight incomplete information, although there is no cooperation in the unique subgame perfect equilibrium of the corresponding complete-information game.

Schmidt (1994) was the first to consider a *non-myopic* opponent; subsequent papers by Aoyagi (1996), Celentani et. al (1994), Evans and Thomas (1997, henceforth ET) all consider only one non-myopic opponent. Let us be precise about the discount

[2]If player 0 were given the opportunity to commit to an action, the one he would choose is called the Stackelberg action. In other words, the Stackelberg action maximises his utility if his opponent plays a best response to whichever action he chooses.

factors $\delta_0, \delta_1, \cdots, \delta_n$ at the cost of some additional notation. All four papers just mentioned deal with the limiting case where all players are patient but 0 is relatively patient, i.e. $\delta_i \to 1 \; \forall i > 0$ and $\frac{1-\delta_0}{1-\delta_i} \to 0 \; \forall i > 0$. One standard justification of a higher discount factor for 0 is that he is a large player who plays with several opponents, who play relatively infrequently. This assumption is not needed by my work, although it is in this context that my work stands in sharp contrast to the previous literature. Here is the main result established by the last mentioned three papers: As the (relatively impatient) opponent also becomes patient in absolute terms, the payoff of the patient player tends towards $g_0^{**} = max\{v_0 \,|(v_0, v_1) \in F^*\}$,[3] where F^* is the feasible and individually rational set of payoffs; i.e. he gets the most that is consistent with individual rationality of the opponent, player 1. Their results thus show that g_0^{**} lower bounds v_{min}^0, the minimum equilibrium payoff of 0 in the incomplete information game as all players become patient(with 0 being more patient).

Now we turn to a key contribution of my paper. The immediate generalisation of the bound obtained by Aoyagi, CFLP, and ET for $n = 1$ to multiple($n > 1$) opponents is $g_0^{**} = max\{v_0 \,|(v_0, v_1, \cdots, v_n) \in F^*\}$. One obvious case where this holds is the one where the more patient player(0) is playing a series of independent games with the other n players.[4] When players $1, ...n$ are myopic, i.e. their discount factors are 0, the above bound follows readily from the analysis of FL; if the other players are also patient the bound derives from (an n-fold repetition) of the analysis of ET. However the previous chapter of this thesis finds that this is not true in general: The presence of additional opponents each with the ability to punish and play out various repeated game strategies complicates the situation. With a single patient opponent[5]

[3] We adopt the convention that the generic value is v; w is the "worst" value; and b denotes the best. The mnemonic advantages hopefully justify any break with practice.

[4] By "independent" I mean that the payoff of any player $i > 0$ is independent of the actions of player $j > 0, j \neq i$. This game is the concatenation of n two-player games, each of which has player 0 as one of the two players. Furthermore, the types of player 0 are independent across these games, in the sense that observing the play in any one game conveys no information about the type of 0 in any other game.

[5] L equals the Stackelberg payoff of 0 in FL.

the previous literature establishes that $v_0^{min} \geq L = g_0^{**}$. It is interesting that, under some non-emptiness restriction that rules out cases like the immediate extension above, a world with multiple opponents gives an *upper bound* on how high a payoff "reputation" can guarantee 0. In the earlier chapter I defined a bound l such that all complete information (i.e. without the perturbed type space) equilibrium payoffs that give player 0 more than l can be sustained in the limiting set of Perfect Bayesian Equilirbrium (PBE) payoffs *even under incomplete information*, and even when I allow arbitarily complex history-dependent types such as those used by Schmidt, Aoyagi, CFLP, and ET; that sequential rationality is satisfied in the construction of the above equilibria makes the task even more onerous. Finally, l could be much lower than g_0^{**}, and even as low as the minmax value of player 0. This means that while reputation is in general valuable to a player, its impact is less dramatic when there are multiple opponents; above a certain payoff for player 0 the perturbed game is no different from the original repeated game of complete information.

Let us now take a closer look at the literature with only one patient opponent. Recall FL's argument for a single *myopic* opponent: If the normal type mimics the type that plays the "Stackelberg action" in every period, then eventually the opponent will play a best-response to the Stackelberg action if she is myopic, thereby assuring 0 of the Stackelberg payoff. By making the opponent patient, Schmidt introduces a twist in the tale: even if the normal type of player 0 perseveres with the Stackelberg action of the stage-game, a non-myopic opponent need not play a period-by-period best response if she fears that she is facing a perverse type that plays like the commitment type on the path but punishes severely if she plays a best response to the commitment type.[6] The source of the problem is that off-path strategies are not learnt in

[6] A related problem is that even if the normal type and a commitment type are the only types, the non-myopic opponent could still fear that play will move to a bad equilibrium of the complete information game after normal type of 0 reveals rationality. If the opponent's best response to the commitment type of 0 does not give the normal type of 0 his best possible payoff, then one may indeed tempt 0 to reveal himself and punish the opponent for best-responding to the commitment type.

any equilibrium with perfect monitoring; this could lead to a very weak lower bound on v_0^{min}, in particular one lower than the Stackelberg bound of FL. Schmidt shows that FL's positive reputation result extends only to "conflicting interest games" — games where the reputation builder wants to commit to an action that minmaxes the other player; in these games the opponent, myopic or not, has no choice about how to respond — she must play her *static* best-response or get less than her minmax value, which is impossible in equilibrium. Later Cripps, Schmidt, and Thomas (CST, 1996) considers arbitrary stage-games and obtains a tight bound that is strictly below the bound of FL. Their bound is "tight" in the sense that there exist equilibria that give 0 payoffs just above this bound.

Subsequent papers argue that CST takes a somewhat bleaker view of reputation effects than warranted. Two of these, Aoyagi and CFLP, both differ from Schmidt and CST in that there is no issue whether strategies can be learnt: Trembles in the first paper and imperfect monitoring in the second ensure that all possible triggers are pressed and latent fears do not remain. When strategies are eventually learnt, reputation is once again a potent force. In particular the lower bound on payoffs is as high as feasibility would allow, and in general the long-run player does even better than getting the static Stackelberg payoff. The final decisive step towards restoring a high lower bound was taken by Evans and Thomas (ET, 1997); they showed that CST gets a weak lower bound because all irrational types are assumed to impose punishments of bounded length; under limiting patience ET extends FL's result to the case of a single long-lived opponent playing an arbitrary simultaneous stage-game under perfect monitoring. Assign a positive prior probability to a type that plays some block of actions, expects a block of replies in return, and punishes for k periods the k^{th} deviation from the desired responses; then the lower bound, with a sufficiently patient opponent, tends towards the *best feasible payoff* for 0 subject to the opponent getting her minmax payoff; thus ET shows that the bound g_0^{**} in Aoyagi and CFLP applies independently of the monitoring structure. In terms of the framework used in the literature, my work adds more opponents to Schmidt's framework.

As mentioned earlier in this section, my work shows that adding more impatient players, who interact with one another, could significantly weaken the impact of reputation. This is proved in the earlier chapter, which makes two important simplifications — it assumes (1) that there is the option to send messages, and (2) that the reputation-builder can commit only to pure strategies of the dynamic game. While each assumption may be natural enough and likely to be satisfied in a range of applications, this chapter seeks to dispense with these and extend the qualitative result. We first note why the assumptions have bite. Both the availability of announcements and the notion of a type as a pure (rather than mixed) strategy of the dynamic game are important for the exact numerical bound and, more substantively, the technique of proof: At each t, the history determines uniquely the pure action that 0 plays if he is indeed the announced type. This enables the other players to mete out a suitably tailored punishment at each time $t+1$ if the normal type of player 0 is trying to mimic a non-normal type; furthermore, any deviation from the announcement is immediately detected by the others. Such tailored punishments and immediate detection may no longer be feasible in either of the following two cases — (i) types are pure strategies of the dynamic game but announcements are not available, or (ii) types could be committed to mixed strategies, either a constant mix independent of history or an arbitrary mixed strategy of the dynamic game. Focus first on the second of the two sources of difficulty pointed out above; suppose we allow for types that commit to mixed strategies of the dynamic game. First, the presence of such types means that the previous bound won't work because at each step there is more freedom for player 0: instead of playing exactly one action from A_0 he can mix among these actions, each such mix inducing a game among the remaining n players. One could still extend the result to mixed strategy behavioural or crazy types when mixing is *ex-post* observable; the bound needs to be redefined to allow for the extra flexibility available to player 0, but the proof remains unchanged in essence —the normal type of 0 cannot get more than the bound if he sticks to his announced mixing probabilities; if he deviates he reveals himself *immediately* as the normal type. However, matters are more complicated when types are committed to mixed strategies that are *unobservable* even ex-post, because deviations are not readily detected. In the latter case the normal type may announce one strategy and follow another that takes the sting out of whatever punishments are intended for him; if the two strategies are observationally indistinguishable— for example if they have the same support on the path— it is impossible to detect deviations on a period-by-period basis. A similar difficulty arises in case (i) above: When there is no scope for messaging, even when the opponents observe a deviation from the equilibrium strategy suggested for the normal type, they cannot learn exactly what type they are facing, as a given deviation may be consistent with a number of different behavioural types. Therefore, they are unable to determine the right punishment. We address these two issues in turn and show that the qualitative message does not rely critically on these.

The plan of this chapter is as follows: section 2 reminds the reader of the model; section 3 retains the assumption that types are pure strategies of the dynamic game, but does not need an announcement phase at the start of the game; the one after allows player 0 to send messages, but broadens the types-space to include types that are committed to mixed/behavioural strategies of the dynamic game; section 5 concludes. The combination of both these extensions — mixed-strategy types without announcements — is left for future research.

2.2 THE MODEL

The model follows the notation of the previous chapter; differences from the baseline model are indicated in the appropriate sections. For the sake of completeness we remind the reader of the notation. There are $n+1$ players — $0, 1, 2, \cdots, n$; throughout we refer to player 0 as "he" and to the others as "she". Player 0 is relatively more patient and attempts to build a reputation. At each time $t = 1, 2, \cdots$ the players play a *simultaneous* stage-game

$$G = \langle N = \{0, 1, \cdots, n\}, \ (A_i)_{i=0}^n \ , \ (g_i)_{i=0}^n \rangle .$$

N is the set of players; A_i is the finite set of pure actions of player i in the stage-game, and $\mathcal{A}_i := \triangle(A_i)$ is the set of mixed actions α_i of player i. An action profile a of all players is in $A := \times_{i=0}^n A_i$; the action vector a_+ of the players $i > 0$ lies in $A_+ := \times_{i>0} A_i$. Throughout, the i^{th} element of a profile/vector is denoted by the same symbol but with the subscript i ; the subscript $+$ denotes all players $i > 0$ collectively. The payoff function of agent $i \geq 0$ is $g_i : A \to \mathbb{R}$; and the vector payoff function is $g = (g_0, g_1, \cdots, g_n) \ : A \to \mathbb{R}^{n+1}$. For any $E \subset \mathbb{R}^d$ and any $J \subset \{1, 2, \cdots, d\}$, the projection of E onto the plane formed by coordinates in J is denoted by E_J:

$$E_J := \left\{ (e_j)_{j \in J} \middle| \exists \, (e_k)_{1 \leq k \leq d, k \notin J} \; s.t. \, (e_l)_{1 \leq l \leq d} \in E \right\}.$$

The convex hull of any subset E of an Euclidean space is coE; it is the smallest convex set containing E. For any player $i \geq 0$ the *minmax value* w_i and the *pure strategy minmax value* w_i^p are defined respectively as

$$w_i := min_{\alpha_{-i}} max_{a_i} g_i(\alpha_{-i}, a_i) \ ; \ w_i^p = min_{a_{-i}} max_{a_i} g_i(a_{-i}, a_i) .$$

The feasible set of payoffs in G is $F := co\{g(a) : a \in A\}$, using an PRD[7](Public Randomisation Device) available to agents; ϕ_t is the observed value of the PRD in period t , and ϕ^t denotes the *vector* of all realised values from period 1 to period t. The individually rational set is $F^* := \{v \in F : \ v_i \geq w_i \ \forall i \geq 0\}$. Monitoring is perfect. This chapter also make the simplifying assumption that mixing by players $1, 2, ..., n$ is observable ex-post.

Players $1, 2, \cdots, n$ all maximise the sum of discounted per-period payoffs; to reduce notation and facilitate comparison with the literature on repeated games I take a common discount factor, i.e. $\delta_i = \delta \ \forall i > 0$.[8] Let us now add incomplete information — the patient player(0) could be one of many types $\omega \in \Omega$; the prior on Ω is given by $\mu \in \triangle(\Omega)$. The type-space Ω contains ω°, the normal type of the repeated game, who maximises the sum of per-period payoffs using his discount factor δ_0. The other types may be represented as expected utility maximisers, although their utility functions may not be sums of discounted stage-game payoffs. Consider, for example, the "strong monopolist" in the chain-store game of Kreps and Wilson; this type has a dominant strategy in the dynamic game to always fight entry. In section 3 below, each type $\omega \neq \omega^\circ$ is identified with or defined by the following sequence

$$\Phi_1(\omega) \in A_0, \Phi_{t+1}(\omega) : A_+^t \to A_0.$$

This sequence $\Phi(\omega) := (\Phi_1(\omega), \Phi_2(\omega), \cdots)$ specifies an initial ($t = 1$) action and for

[7]Given any payoff v in the convex hull of pure action payoffs, Sorin constucted a pure strategy *without public randomisation* alternating over the extreme points so as to achieve exactly v when players are patient enough. This does not immediately allow us to get rid of the PRD because the construction of Sorin need not satisfy individual rationality — after some histories the continuation payoff could be below the IR level. Fudenberg and Maskin extended his arguments and showed that this could be done so that after any history all continuation payoffs lie arbitrarily close to v ; if v is strictly individually rational so are the continuation values lying close enough. Taken together these papers showed that a PRD is without loss of generality when players are patient. I continue to make this assumtion in the interests of expositional clarity.

[8]Lehrer and Pauzner (1999) look at repeated games with differential discount factors, and find that the possibility of temporal trade expands the feasible set beyond the feasible set of the stage-game. This creates no problems for my results because the feasible set of the stage-game continues to remain feasible even if all δ_i's are not equal .

each $t > 1$ maps any history of actions[9] played by players $1, 2, \cdots, n$ into an action of player 0. In section 4 we expand the potential set of types to allow the reputation-builder to commit to mixed strategies of the dynamic game: Each type $\omega \neq \omega^\circ$ will then be defined by the following sequence

$$\Phi_1(\omega) \in \triangle(A_0), \Phi_t(\omega) : A^{t-1} \to \triangle(A_0).$$

Fix (G, Ω, μ), where G is the stage-game, $\Omega = \{\omega^\circ, \omega_1, \cdots, \omega_K\}$ is an arbitrary finite set of types, and μ is the prior on Ω. In order to give reputation effects room to operate I allow a very rich set of types; in particular any of the complex history-dependent behavioural types employed by CFLP or ET may be included in Ω. The limiting payoff set of the complete information game is V^{ci};V is the limit of the PBE payoffs of the incomplete information game as $\delta_i \to 1$ and $\frac{1-\delta_0}{1-\delta_i} \to 0$ for all $i > 0$; and v_0^{min} denotes the infimum of 0's payoffs in V. My results are of the form $v_0^{min} \leq l$.

2.3 REPUTATION WITHOUT ANNOUNCEMENTS

The previous chapter of this thesis poses the problem as a signalling game, where 0 is first given the opportunity to state his type. While adding an announcement phase is a natural modification, one might still wonder if the result goes through without announcements.[10] The earlier chapter implicity assumes that a "crazy" type does not misrepresent its type, although a normal type may choose to mimic a "crazy" type. Truth-telling is incentive compatible for the "crazy" types, but only weakly so. If we modify the result to work even without an announcement stage, this assumption is

[9]Potentially player 0 could condition his play on the PRD upto period $t+1$, i.e. on ϕ^{t+1} , in addition to a_+^t. This would not change our results or proofs; the notation is more involved though.

[10]The results on the previous section titled "Long-run Player Moves First" goes through even without annoucements, which are irelevant when 0's action is observed before the normal players 1,2,...n move. It is interesting to compare this with FL, where the result is immediate if the patient player moves first.

little more than a technical convenience. Fortunately one can indeed show that the main result extends, although at the cost of added complexity of the strategies.[11]

At each $t = 1, 2, \cdots, \infty$ players $i \geq 0$ simultaneously choose actions $a_i \in A_i$. The public history of actions upto and including time t is $h^t \in A^t$. An equilibrium comprises the following elements, defined recursively over the time index:

(i) for type $\omega \neq \omega^\circ$ of 0 a period-1 action $\sigma^\omega(1) \in A_0$; and a sequence of maps

$$\sigma^\omega(t) \ : \ A_+^{t-1} \to A_0; \ t > 1 \ ;$$

(ii) for each $i \geq 0$ a map $\sigma_i(1) \ : \ \{\phi^1\} \to A_i$, and for each $t > 1$ maps

$$\sigma_i(t) \ : \ A^{t-1} \times \{\phi^t\} \to A_i \ ;$$

(iii) beliefs $\mu\left(. \left| h^{t-1}, \sigma\right.\right) \in \triangle(\Omega)$, following the history h^{t-1} and given the equilibrium strategy profile σ, are obtained by updating using Bayes' rule wherever possible. Additionally we stipulate that for $\omega \neq \omega^o$ the strategy σ^ω is given by $\Phi(\omega)$, as long as ω has not violated his own precepts. This in effect defines a "crazy" or "behavioural" type. Let $\sigma^\omega := \{\sigma^\omega(t)\}_{t \geq 1}$. The strategy σ_i of player $i \geq 0$ is the collection of maps $\{\sigma_i(t)\}_{t \geq 1}$; the set of all strategies of player i is Σ_i. As usual, a strategy profile is

$$\sigma := (\sigma_0, \sigma_1, \cdots, \sigma_n) \in \Sigma := \Sigma_0 \times \Sigma_1 \times \cdots \times \Sigma_n.$$

In what follows $u_i(\bullet)$ refers to the discounted value of a strategy profile to player i, possibly contingent on a certain history of actions; thus $u_i\left(\sigma \left| h^{t-1}\right.\right)$ is the sum of the per-period payoffs of player i discounted to the beginning of period t, following the history h^{t-1}, and given that players play according to the strategy profile σ. We refer

[11] The *critical* value of the discount factor required for the proof to work would in general be dependent on the cardinality of the type-space, but the limiting value is independent of the fine details.

to the payoff of the normal type of player 0 when we use $u_0(\bullet)$. Our equilibrium concept is stated below.

Definition : A tuple $\left(\sigma_0^*, \sigma_1^*, \cdots, \sigma_n^*, \{\mu\left(. \left|h^{t-1}, \sigma^*\right.\right)\}_{t>1}\right)$ defines a Perfect Bayesian Equilibrium (PBE) if no player has a strictly profitable unilateral deviation, i.e. for any $i \geq 0$ and any $(t-1)$-period history h^{t-1} we have

$$u_i\left(\sigma^* \left|h^{t-1}\right.\right) \geq u_i\left(\hat{\sigma_i}, \sigma_{-i}^* \left|h^{t-1}\right.\right) \; \forall \; \hat{\sigma}_i \in \Sigma_i.$$

First, a few comments on the intuition and the strategy of proof of Proposition 1 below, which extends Proposition 1 of the previous chapter to the current scenario where announcements are unavailable. Consider a strategy σ^{ci} of players $0, 1, \cdots, n$ in the repeated game giving player 0 a payoff of $l + \epsilon$ and any $i > 0$ at least W_i. We shall now construct an equilibrium of the incomplete information game in which ω° plays according to σ_0^{ci}. Consider the following strategy profile. Players start by playing according to σ^{ci}; as soon as 0 deviates from σ_0^{ci} the prior μ is appropriately modified. Find the behavioural type ω^* of player 0 who gets the highest probability in the updated beliefs. The equilibrium strategy asks players $i > 0$ to play $\rho\left(\Phi(\omega^*)\right)$the next period. Of course it is possible that the actual action played by 0 in the next period is something other than $\Phi(\omega^*)$. It is important to note that if $|supp(\mu)| = K$, then along any path there can be at most K points where all types that are currently in the support do not behave identically; this is the content of Lemma 1. At each such point at least one type is eliminated from the support of the posterior derived from μ by Bayesian updating. Thus there can be no more than K mistakes in predicting the strategy of 0. Consequently in at most K periods can the normal type of player 0 get more than $\hat{v}_0 := l + \epsilon$. More than K mistakes is a 0-probability event under the equilibrium strategy and leads players $1, 2..., n$ to (rightly) believe that they are facing the normal type who has deviated. Although this means that the normal type of player 0 cannot gain by deviating at more than K points, it is not enough for my proof to go through. There are two potential problems with this line of argument. First, since each deviation gives player 0 a one-shot gain, it might be tempting for a normal-0 to deviate at points of time that are very far apart. It might take so long for his true type to be discovered that punishments could become irrelevant from the point of view of the current period. Since he can choose any strategy he wishes, there is nothing to prevent him from deviating at points of time very far apart. To prevent this we apply a "real-time" correction. When a mistake has been made in predicting the strategy of 0, the others are asked to apply an *immediate* correction by minmaxing 0 to wipe out the gains that a normal type of 0 would have accumulated by that mistake. This need not be viewed as a punishment but as a corrective. The second potential problem is : Why would $i > 0$ stick to her assigned role? The argument combines the proof of Proposition 1 of the preceeding chapter with the observation that if i deviates she can escape punishment, and perhaps even get high payoffs, for at most K periods; if δ is high enough, then punishments cannot be avoided for long enough— conditionally minmaxing her is enough. Before we state Lemma 1 formally, we recall the following — $\mu\left(\bullet|h^{t-1},\sigma\right)$ denotes the posterior at the start of the period t derived from the prior μ, conditional on the history of play upto period $(t-1)$ and the equilibrium strategy σ; for convenience we sometimes denote it by μ_{t-1}.

LEMMA 1 (Finite Non-identical Play Property): *Suppose $\omega^\circ \notin supp\left(\mu_T\right)$, for some $T > 0$. Along any path of play, equilibrium or otherwise, there can be at most $|\Omega - \{\omega^\circ\}| = K - 1$ periods after T where all types of player 0 that are still in the support of the posterior beliefs do not play the same action.*

Proof: Let $A\left(\mu_{t-1}\right)$ denote all actions that have positive probability of being played in period t by player 0 given μ_{t-1}. Suppose that $suppA\left(\mu_{t-1}\right) > 1$. If $a_0(t) \notin suppA\left(\mu_{t-1}\right)$, then Bayes' rule imposes no restrictions on how μ_t is to be derived from μ_{t-1}; so we set $\mu_t(\omega^\circ) = 1$. Otherwise, if $a_0(t) \in suppA\,(\mu_{t-1})$, we must have $supp\,(\mu_t) \leq supp\,(\mu_{t-1}) - 1$, since all types in the support of $\mu(t-1)$ that were expected to play the action a_0^t may now be dropped from the support of $\mu(t)$. For any path of play where $suppA\,(\mu_t) > 1$ for $t = t_1 < t_2 < ... < t_K$, we must have $suppA\,(\mu_s) = 1 \; \forall s > t_K$. ■

Proposition 1 below formalises the argument in the above paragraph using lemma 1(FNPP). The proof continues to employ the simplifying assumption that mixing by players $i > 0$ is observable ex-post; as we have seen in the previous chapter, relaxing this is fairly uncomplicated.

PROPOSITION 1 (Upper Bound Without Announcements) *Under assumptions N and FD, $v_0^{min} \leq l$ even if announcements are unavailable.*

***Proof*:** Fix $\epsilon > 0$. We constuct a PBE of the incomplete information game in which player 0 gets $\hat{v}_0 = l + \epsilon$ and any $i > 0$ gets at least W_i. Using TPL from the previous chapter choose $\eta \in \mathbb{R}^n$,$\Delta > 0$ such that

$$W_i < \eta_i < \eta_i + \Delta < \lambda_i \; \forall i > 0.$$

As before, λ_i is the maximum payoff of player i across all slices when in each slice players $i > 0$ play accoding to ρ. Define the length P of the punishment phase such that

$$PW_i + K \times max_{a \in A}\; g_i(a) < P\eta_i + K \times min_{a \in A}\; g_i(a) \; \forall i > 0 \cdots (*)$$

In what follows a "deviation" by player 0 in period t from a strategy profile σ is taken to mean that in period t player 0 plays an action $a_0(t) \notin A\,(\mu(.\,|h^{t-1}\,,\sigma))$. Beliefs are updated using Bayes' rule following an event that has positive probability; following a 0-probability event under σ, the posterior puts probability 1 on the normal type of player 0 as in lemma 1 above. Let $\sigma^{ci}(v_0, v_i, \cdots, v_n)$ denote the equilibrium(SPNE) strategy profile of the complete information repeated game that gives ω° a payoff of v_0, and gives player $i > 0$ a payoff of v_i; when arguments are omitted it means that we do not impose any restrictions on those payoffs except that they exceed the corresponding minmax levels.

Let $\omega^*(t)$ be the type on which μ_{t-1} places the highest probability; if more than one type has this property choose the one with the lowest index when Ω is enumerated in any fixed way. (Indeed we could just as well choose this using any other deterministic algorithm.) Now let $a_o^*(t) := \Phi_t(\omega^*(t))$. Note that if $A(\mu_{t-1})$ contains only one action, then this action is $a_o^*(t)$. Then define

$$\Psi_t(\mu_{t-1}) := \rho(a_o^*(t)) = \rho(\Phi_t(\omega^*(t))) \in \mathcal{A}_+,$$

the n-tuple of actions that delivers $\hat{v}_0 = l + \epsilon$ in slice-$a_o^*(t)$. We assert that the following strategies are part of a PBE giving ω^0 a payoff of $\hat{v}_0 := l + \epsilon$. Play proceeds according to the following phases.

Phase I: In period t play the action n-tuple $\Psi_t(\mu_{t-1}) \in \mathcal{A}_+$. If $a_o^*(t) = a_0(t)$, i.e. if the other players correctly anticipated the action of player 0, continue in $PhaseI$. If $a_o^*(t) \neq a_0(t)$ but $a_0(t) \in supp(\mu_{t-1})$, i.e. player 0 played an action that was in the support but not the action $a_o^*(t)$ that the others had bargained for, switch to *Phase Adjust* : Minmax 0 for P periods, and return to *Phase I* if there are no deviations by players $i > 0$.

Now we specify how to react to deviations. Suppose player 0 has never deviated(i.e. all actions played by 0 have been consistent with some type of 0, normal or not), and that player $i > 0$ deviates at time τ ; then play enters *Phase* $II(i)$, followed by *Phase* $III(i)$.

Phase $II(i)$: Player i is conditionally minmaxed in slice $a_o^*(t)\left(h_+^{t-1}\right)$ at time $t =$

$\tau+1,\cdots,\tau+P$.

Phase $III(i)$: Give the expected payoff vector $\eta(i)$ to players $i>0$ at $t=\tau+P+1,\tau+P+2,\cdots,\infty$. The i^{th} component of $\eta(i)$ is η_i, and all other components $j>0$ are $\eta_j+\Delta$; $\eta(i)$ incorporates a small reward of Δ for each player $j>0$ other than the last player(i) to deviate.

If player 0 deviates during any phase— including the case $a_0(t)\notin supp\left(\mu_{t-1}\right)$— , we set $\mu\left(\omega^{\circ}\left|\omega,h^{\tau}\right.\right)=1$; player 0 is then minmaxed to wipe out gains, followed by a switch to $\sigma^{ci}\left(\lambda_0\left(\underline{a}_0\right),\lambda_1\left(\underline{a}_0\right),\cdots,\lambda\left(\underline{a}_0\right)\right)$, where $\underline{a}_0\in argmin\left\{\lambda_0\left(a_0\right)\right\}$.

This may be verified to be a sequentially rational equilibrium, in a manner similar to Proposition 1 of the previous chapter. ■

2.4 REPUTATION WITH TYPES THAT MIX

All previous results are proved for a type space $\Omega=\{\omega^{\circ},\omega_1,\cdots,\omega_K\}$, where the behavioural types $\omega_1,\cdots,\omega_K$ are committed to *pure* dynamic game strategies, i.e. each type $\omega\in\Omega-\{\omega^{\circ}\}$ plays in each period $t=1,2,...$ a unique action $\Phi_t(\omega)\left(h_+^{t-1}\right)\in A_0$ conditional on the history h_+^{t-1} of action profiles of players $1,2,\cdots,n$. Each type $\omega\neq\omega^{\circ}$ is thus identified with or defined by the following sequence of mappings

$$\Phi_1(\omega)\in A_0,\Phi_t(\omega):A_+^{t-1}\to A_0$$

, which specifies an initial action (for $t=1$) and for each $t>1$ maps any history of actions played by players $1,2,\cdots,n$ into an action of player 0. This notion of a type as a pure (rather than mixed) strategy of the dynamic game is important both for the exact numerical bound and, more substantively, the technique of proof: At each t, given the history h^{t-1}, the pure action $\Phi_t(\omega)\left(h_+^{t-1}\right)$ that 0 plays is known to the other players $1,2,\cdots,n$; this enables them to not only mete out a suitably tailored punishment, but also detect deviations by the reputation builder as soon as they occur. For the case of mixed strategy crazy types we maintain the earlier notation — fix (G, Ω, μ), where G is the stage-game, $\Omega = \{\omega^\circ, \omega_1, \cdots, \omega_K\}$ is an arbitrary finite set of types, and μ is the prior on Ω. Each type $\omega \neq \omega^\circ$ is now identified with or defined by the following sequence

$$\Phi_1(\omega) \in \triangle(A_0), \Phi_t(\omega) : A^{t-1} \to \triangle(A_0)$$

,where each A_0 has been replaced by $\triangle(A_0)$, and and type ω's period-t strategy maps from A^{t-1} rather than A_+^{t-1}. For a behavioural type it is unnecessary to specify how he plays after having himself deviated in the past, because this is inconsistent with the definition of a behavioural type as one that is committed to playing in a certain way. When we consider pure strategy types, all the relevant information is summarised in the history h_+^t of the opponents' moves; with mixed strategy types several actions may be consistent with any given type, and we can specify how 0 plays conditional on his earlier *realised* actions. Therefore the domain of $\Phi_t(\omega)$ is A^{t-1} rather than A_+^{t-1}.

Using relatively standard techniques I extend my results in proposition 3 of the chapter before to the case where mixing by players $1, \cdots, n$ is not observable ex-post; however allowing the reputation-builder to commit to mixed strategies requires significantly different techniques. Before a formal statement of the proposition let me try to explain the key issues involved, starting with the conceptually simpler case when $0's$ mixed strategies are ex-post observable. First, the presence of types that mix means that the previous bound won't work because at each step there is more freedom for player 0: instead of playing exactly one action from A_0 he can mix among these actions, each such mix inducing a game among the remaining n players. We must redefine the upper bound to allow for the extra flexibility available to player 0, but the proof remains unchanged in essence — if he deviates he reveals himself *immediately* to be a normal type; micmicking another type faithfully isn't a profitable deviation from playing like a normal type. Matters are more complicated when types are committed to mixed strategies that are *unobservable* even ex-post, because deviations are not readily detected. This point is explained in greater depth further in the section.

The notation below closely follows that for pure types. Here is the concept of a *slice* when there are mixed-strategy types:

Definition **:** For any $\alpha_0 \in \mathcal{A}_0$, the *slice*-$\alpha_0$ is formally the game $G(\alpha_0)$ induced from G by replacing A_0 by $\{\alpha_0\}$, and restricting the domain of g_i to $\{\alpha_0\} \times A_+$ to get $\hat{g}_i$, i.e.

$$G(a_0) := \left\langle N := \{0, 1, \cdots, n\}; \{a_0\}, (A_i)_{i>0}; (\hat{g}_i)_{i\geq 0} \right\rangle.$$

Our earlier upper bound was defined by taking the max of $w_0(a_0)$ over all actions $a_0 \in A_0$. Now we need to take the *supremum* over all probability distributions α_0 on A_0. This is clearly a larger set and consequently the upper bound cannot decrease. We now redefine the bound l. First, for each mixed action $\alpha_0 \in \triangle(A_0)$ we have an induced game $G(\alpha_0)$ as before.

Definition **:** The *conditional minmax* of $i > 0$ in the slice-α_0 is the minmax of $i > 0$ *conditional* on the slice $G(\alpha_0)$; call it $w_i(\alpha_0)$.

Definition **:** The *maxminmax* of a player $i > 0$ with respect to 0 is defined as the supremum of all conditional minmaxes:

$$W_i := sup_{\alpha_0} w_i(\alpha_0).$$

Start by defining

$$\mathcal{F}(\alpha_0) \ := \{v \in F(\alpha_0) \ : \ v_i \geq W_i \forall\ i > 0\}.$$

Note that $\mathcal{F}(\alpha_0) \subset \mathbb{R}^{n+1}$, the projection of $\mathcal{F}(\alpha_0)$ onto the 0^{th} coordinate is denoted by adding the subscript 0 , i.e. $\mathcal{F}_0(\alpha_0) \subset \mathbb{R}$. First, observe that the worst payoff for player-0 in the slice-α_0 subject to each player $i > 0$ getting above her maxminmax is

$$w_0(\alpha_0) := inf\ \mathcal{F}_0(\alpha_0) \equiv inf\{\ v_0 :\ \ (v_0, v_+) \in \mathcal{F}(\alpha_0)\}.$$

The new upper bound is given by the supremum of these infima over all slices :

$$l\ := sup_{\alpha_0} w_0(\alpha_0) \equiv sup_{\alpha_0} inf\ \mathcal{F}_0(\alpha_0).$$

Even with this new bound, designing and enforcing a punishment for announcing a non-normal type proves challenging. To avoid further complications at this point, we retain the announcement phase. Since announcements are available we allow player 0 to send an initial message m and subsequent messages m_t at each time $t = 1, 2, \cdots$. We also make the simplifying assumption that a Random Public Device (PRD) is available. For simplicity assume that mixing by players $i > 0$ is observable ex-post, even though we cannot observe mixing probabilities chosen by 0; this simplifies the proof considerably without affecting the essence of the result.[12]

With types that mix the crux is that detecting deviations is no longer immediate — ω° can announce some $\omega \neq \omega^\circ$ but behave like $\sigma^\sim \neq \sigma(\omega)$; there is no way to immediately detect a discrepancy between the initially announced type and the actual play if, for example, $\sigma(\omega)$ and $\sigma^\sim$ have the same support in A_0 at each t on the path of play. The key to resolving this problem is that while there is no way of detecting a deviation period by period, one may monitor the *distribution* of the actions of 0 over time. More precisely, given any announced ω we can define for each

[12]In other words, even if we assumed that mixing by all players $0, 1, ..., n$ is unobservable even ex-post we can define and prove a non-trivial upper bound on reputation. We would then use the minmax (w_0^p) of 0 over A_+ rather than $\triangle A_+$ during the *adjustment phase*, to be defined shortly. The new bound would merely be $max\{w_0^p, l\}$,where l is the bound defined above. In particular the above argument works unchanged when (a) $w_0^p = w_0$, or (b) $w_0^p \leq l$, where (a) implies (b). Since $l < g_0^{**}$, so is $max\{w_0^p, l\}$, ensuring that the bound will be non-trivial.

period t an *excess reward function* ξ_t, which measures the excess of the *realised* payoff of 0 over the expected payoff. Over time a Law of Large Numbers(henceforth LLN) would, we might hope, assure us that player 0 won't be too much above his expected payoff if he is playing according to the initially announced type.

At a technical level, one should note that classical LLNs use independent and identically distributed random variables. The requirement that the random variables ξ_t be identical is easily relaxed since they are uniformly bounded in variance. (This follows immediately from the payoff functions $g_i : A \to R$ being uniformly bounded by $\pm M$, where $M := max_i max_a |g_i(a)|$; it follows that $|\xi_t| \leq 2M \; \forall t$, which together with $E\xi_t = 0$ implies that $V\xi_t \leq 4M^2$.)

The first condition (independence) is not unfortunately not met because types condition their t period play on the history of actions upto $t-1$; in other words, a type can play a different mix at different histories. To avoid this problem we could assume that the type space Ω contains only constant (mixed) action types, i.e. each $\omega \neq \omega^\circ$ plays $\alpha(\omega) \in \triangle A_0$ after all histories. Then players $i > 0$ also need to play a constant profile $\rho(\alpha(\omega))$ to punish 0 for declaring a non-normal type ω. Since the mixed action profile is history-independent as long as there are no deviations, ξ_t's are indeed IID; this case is readily covered by LLNs. However reputational arguments derive considerable power from the reputation builder's ability to condition his future actions on the past play of his opponents.

Fortunately, even when one allows such permissive history-dependent behavioural types that mix we can show that our construction admits of an appropriate LLN (for dependent random variables) is applicable. We now formally define the excess payoff functions and state the appropriate LLN.

Definition **:** The excess payoff function $\xi^\alpha : A \to R$ is defined by $\xi^\alpha(a) := g_0(a) - g_0(\alpha)$, where α is any mixed action profile.

This is the excess payoff to player 0 when the action profile played is a and the desired mixed action was α. Using this we next define the excess payoff *given* any announced type ω and any $t-1$ period history h^{t-1} when players follow the punitive plan encoded in $\rho(\circ)$.

Definition : $\xi_t\left(\omega, h^{t-1}\right)(a) := \xi^{\beta}(a)$, where $\beta = \left(\Phi_t(\omega)\left(h^{t-1}\right), \rho\left(\Phi_t(\omega)\left(h^{t-1}\right)\right)\right)$

Before proceeding further we state some results that will enable us to state an appropriate weak law of large numbers. Let $Y_1, Y_2, \cdots$ be a sequence of zero-mean random variables , and let I_{t-1} is the information filtration at time $t-1$. Then we define the following:

Definiton : $\{Y_t\}_{t\geq 1}$ is a Martingale Difference Sequence if $E\left(Y_t\left|I_{t-1}\right.\right) = 0 \forall t$.

The next property records that the random variables ξ_t form a MDS, where I_{t-1} now denotes the filtration generated by $\xi_{t-1}, ..., \xi_1$:

Property : $E\left(\xi_t\left|I_{t-1}\right.\right) = 0$ and $V\xi_t \leq 4M^2 \ \forall \omega, h^{t-1}$.

Either of the following theorems may, for example, be used (for the first see Badi Baltagi, pg. 218).

THEOREM 1: Let $\{Y_t\}_{t\geq 1}$ be a MDS. If $\sum_{t=1}^{T} \frac{\sigma_t^2}{t^2} < \infty \forall T$ then $\frac{1}{T}\sum_{t=1}^{T} Y_t \longmapsto^{prob} 0$.
Fact 1: $\sum_{t=1}^{\infty} \frac{1}{t^2}$ is a convergent series.

***THEOREM 2 (Chow):*:** Let $\{Y_t\}_{t\geq 1}$ be an MDS w.r.t $\{I_t\}_{t\geq 1}$. If for some $r \geq 1$

, $\sum_{t=1}^{T} \left(E \left| Y_t \right|^{2r} \right) / t^{r+1} < \infty$, then $\frac{1}{T} \sum_{t=1}^{T} Y_t \longmapsto^{a.s.} 0$.

Fact 2 **:** Convergence almost surely implies convergence in probability.

LEMMA 2**:** Along any equilibrium path and any ω, we have $\frac{1}{T} \sum_{t=1}^{T} \xi_t \longmapsto^{prob} 0$.

PROOF**:** The proof follows from the properties of the excess payoff functions and either (a) Theorem 1 and Fact 1, or (b) Chow's theorem, Fact 1, and Fact 2 when we take $r = 1$. ■

LEMMA 3 **:** For any ω and any history h^t , $\exists$ N and $\bar{\delta_0} \in (0, 1)$ such that if $\delta_0 > \bar{\delta_0}$, the total discounted sum of excess payoffs

$$\sum_{t=1}^{T} \frac{\xi_t + \delta_0 \xi_{t+1} + \cdots + \delta_0^{N-1} \xi_{t+N-1}}{N} < \epsilon.$$

Before an exact statement of the proposition and its proof, let me present a sketch highlighting the key features of the construction. Recall the structure of announcements — player 0 sends an initial message $m \in \Omega$ at $t = 0$, and at each time $t \geq 1$ a message $m_t \in \Omega$ just *before* the stage-game G is played. (Thus each t is a length rather than instant of time.) Beliefs are formed in the natural way — in keeping with Bayes' rule players $i > 0$ put probability 1 on the announced type m; following any unexpected message or play, players put probability 1 on the normal type. As before the normal type is asked to send the message ω° at all t. Any type $\omega \neq \omega^{\circ}$ sends the message ω at all t. If $m = \omega^{\circ}$, then play moves to the SPNE of the complete information game (see Fudenberg and Maskin (1986) giving player 0 a payoff of $v_0 = l + 4\epsilon$. If $m = \omega \neq \omega^{\circ}$, then play starts in Phase Ia, where for N periods play is according to $\rho(\circ)$, which tracks the (mixed) strategy of player 0 and punishes him on each slice. By the definition of $\rho(.)$ the expected payoff of 0 at each time t

does not exceed $l + \epsilon$. However even if 0 mixes according to the initially announced type, there is a positive probability that the discounted sum of *actual* excess payoff during any $PhaseIa$ exceeds $l + \epsilon$. In order to ensure that these excess payoffs to not accrue, we need to do a review after the N periods of $Phase\ Ia$,— which was not necessary when $\Phi_t(m)\,(h^{t-1})$ was a degenerate distribution over the actions in A_0. If the realised payoff of 0 during the block of N periods exceeds the target payoff $(= l + \epsilon)$ by more than ϵ the other players mixmax 0 for exactly NP periods so as to wipe out any gain over and above the permissible limit of ϵ. (If δ is high then players $i > 0$ would rather carry out a costly punishment for NP periods than be punished into the future.) At the end of this adjustment phase we move back to the start of $Phase\ Ia$ for another cycle of N periods. This ensures that if we target $l + \epsilon$, then 0 cannot get more than $l + 2\epsilon$. With types that mix it is thus a positive probability event that the adjustment phase is triggered even *on* the equilibrium path; we need to ensure that this is not often enough to reduce the expected utility from the equilibrium path to where players $1, \cdots, n$ would rather face punishment than carry out this expensive threat against player 0. However if we choose N to be a large enough integer this probability can be made arbitrarily small, say θ; lemma 2, together with the fact that players are patient, assures us of this. The proof makes the argument exact by computing bounds on the payoffs of players $i > 0$ during the on-path play, and verifies that appropriate incentives are in place. It would be amiss to not mention the pioneering work of Radner, whose "review stategies" this resembles. In the repeated principal-agent model Radner(1985) shows that one can get approximately effcient subgame perfect Nash equilibria by using a very similar form of statistical testing followed by a reversion to the stage-game that is Pareto-worse compared to the suggested point. The critical difference is that in Radner's work it is enough to monitor the realisation of successes and failures generated by the (unobservable) actions of the agent and then use a LLN for IID random variables; this is not so for my proof for a couple of reasons. First, note that a WLLN for independent trials is applicable in our case only when player 0 plays a constant mixed action in each period; otherwise a LLN for dependent random variables is called for. Secondly, unlike in Radner's model, the variable being monitored is generated by the actions of not just the player who is being monitored but by others as well — so we need to give appropriate incentives so that the others play actions that allow us to keep tabs on 0's actions statistically. In other words, the outcomes (success or failure) are independent over periods in Radner's case, but in mine I need the other players to choose their actions in a (complicated) history-dependent way so as to ensure that the excess rewards form a martingale difference sequence.

Since players $i > 0$ carry out this strategy the normal player 0 cannot get more than $v_0 - 2\epsilon$ by declaring $m = m_t = \omega \neq \omega^\circ$ and following $\sigma^\sim$; this is strictly worse than the payoff to declaring $m = m_t = \omega^\circ$ and following the SPNE $\sigma^{ci}(v_0)$ of the complete information game. What does the strategy ask ω° to do after declaring $m \neq \omega^\circ$? He is then asked to reveal himself at the first possible opportunity. For example, if type ω° has until time τ mimicked ω then he is asked to send the message $m_{\tau+1} = \omega^\circ$ and reveal himself ; if τ belongs to a Phase Ia that started at $s < \tau$ then play switches to a game of complete information that gives player-0 a payoff of $v_0 - \epsilon = l + 3\epsilon$ from s onwards. (This can be done in such a way that others get payoffs above their respective w_i. Note that in order to give 0 a payoff of $l + 3\epsilon$ from the *start* of the current phase Ia players $i > 0$ might need to give him more than $l+3\epsilon$ in the continuation game when he deviates. If all players are patient this will cost them no more than a small amount π, which would still be above their maxminmax. Note: Indeed all we need is to give $i > 0$ a payoff above the minmax w_i, which is in general lower than the maxminmax W_i.) The following proposition is the analogue of proposition 1 of chapter 2 for types spaces containing behavioural types that mix when mixed strategies (and mixing is not ex-post observable).

PROPOSITION 2: *Under assumptions N and* FD *(with announcements), we have* $v_0^{min} \leq l$ *even if the type space contains behavioural types committed to strategies that involve mixed and/or have infinite memory.*

***PROOF*:** Fix an arbitrarily small $\epsilon > 0$. (Note: As we shall see readily ϵ cannot be so large that the target payoffs are infeasibly high.) It is sufficient to find a PBE in which ω^0 gets an equilibrium payoff of $v_0 := l + 4\epsilon$. In what follows let $\sigma^{ci}(x_0)$ denote a SPNE for the complete information game that gives player 0 a payoff of x_0. First we describe the strategy profile that gives ω° a payoff of v_0 in the reputational game. We assume for convenience that a PRD is available.

Description of the Strategy Profile :

The normal type is asked to send the message ω° at all t, and play according to $\sigma^{ci}(v_0)$. Conditional on having announced a non-normal type until date τ, ω° is expected to reveal himself at the start of $\tau + 1$ by sending the message ω°. (Any type $\omega \neq \omega^\circ$ sends the message ω at all t and plays according to $\Phi(\omega)$.) The path of play for players $i > 0$ depends on the initial announcement, as detailed below.

Case I $(m = \omega^\circ)$*:* If $m = \omega^\circ$, then play the strategy profile $\sigma^{ci}(v_0)$ starting at time $t = 0$.

Case II $(m \neq \omega^\circ)$*:* If $m \neq \omega^\circ$ and no player has ever deviated, then we are in $Phase\ I$. In what follows we use a PRD; therefore "deviation" by $i > 0$ refers to deviation from a *pure* action contingent on the realised draw; for player 0, a "deviation" is said to occur at period t only when an action $a_0(t)$ not in the support of $\Phi_t(m)(h^{t-1})$ is played.

We first define the following quantities. Lemma 4 assures us that with mixed types there is an "appropriate" punishment in each slice:

LEMMA 4: *Fix any slice* $\alpha_0 \in \triangle A_0$. $\exists \rho(\alpha_0) \in \Delta A_+$ *such that*

$$g_0(\alpha_0, \rho(\alpha_0)) \leq l + \epsilon \text{ and } \lambda_i(\alpha_0) := g_i(\alpha_0, \rho(\alpha_0)) > W_i + \gamma \text{ for some } \gamma > 0,$$

where γ *depends on* ϵ.

PROOF : similar to corresponding lemma of the previous chapter.

This implies that $\lambda_i := sup_{\alpha_0} \lambda_i(\alpha_0) > W_i + \gamma$. Now fix the $n+2$ quantities $\pi, (\eta_i)_{i>0}, \Delta$ s.t. $W_i < \eta_i < \eta_i + \Delta < \lambda_i - \pi < \lambda_i$. Since $W_i < \eta_i$ by construction, we can find a large enough interger P , which will later be seen to be the length of the conditional minmaxing phase, such that

$$PW_i + max\ g_i(a) < P\eta_i + min\ g_i(a)\ \forall i > 0 ....(*)$$

We start in $Phase\ Ia$: During the first N periods, where N is chosen to satisfy lemma 3 above, and the prescribed play at t conditional on the $t-1$-period history h^{t-1} is $\rho(\Phi_t(m)(h^{t-1}))$, the mixed action profile that gives player 0 a low expected payoff(no more than $l + \epsilon$) in $\Phi_t(m)(h^{t-1})$. After N periods we now do a review, which was not necessary when $\Phi_t(m)(h^{t-1})$ was a degenerate distribution over A_0. If the realised payoff of 0 during the block of N periods does not exceed the target payoff by more than ϵ, we start another Phase Ia; otherwise the other players mixmax 0 for NP periods, wiping out any gain over and above the permissible limit of ϵ, while the normal type of 0 is instructed to play some pure best reply during this stage. At the end of this adjustment phase IIb we move back to the start of phase Ia for another cycle of N periods. By lemma 3 we can find a large enough integer N which is independent of the discount factors and the type ω such that the probability of the average discounted excess payoff being more than 2ϵ after N periods falls short of θ, whre $\theta > 0$ will be chosen to be suitably small later on.

Suppose player 0 has never deviated from his announced strategy in the past and that player $i > 0$ deviates from the prescribed path at time τ ; then play enters *Phase* II_i , where player i is conditionally minmaxed in slice $\Phi_t(m)\left(h^{t-1}\right)$ at time $t = \tau, \tau+1, \cdots, \tau+P$, where P is the length of the punishment. Once *Phase* II_i is over play moves into *Phase* III_i , where the play is such that $\eta(i)$ is delivered at $t = \tau+P+1, \tau+P+2, \cdots, \infty$; recall that the i^{th} component of this vector is η_i , and all other components $j > 0$ are $\eta_j + \Delta$. If player j unilaterally deviates from *Phase* II_i or *Phase* III_i, then impose *Phase* II_j followed by *Phase* III_j, and so on.

Conditional on an announcement $m = \omega \neq \omega^{\circ}$, players believe that they are indeed facing type ω until 0 deviates from $\Phi(m)$ or sends a message $m \neq \omega$ and reveals himself to be the normal type. Once the continuation game is a complete information game, any player including 0 may be punished by minmaxing by the other players P periods (see * above). If player 0 reveals himself as the normal type at the start of a *Phase Ia* then play re-adjusts to give him $v_0 - \epsilon = l + 3\epsilon$ from the start of the current *Phase Ia.* If he reveals normality during a *Phase Ib* then the minmaxing continues until the current phase is played out, and thereafter play switches to $\sigma^{ci}(v_0 - \epsilon)$.

Checking that the Strategy Profile Constitutes a PBE:

Step 1 : Player 0 will not deviate :

(i) From standard results for the complete information game it follows that 0 does not deviate after he reveals himself to be a normal type.

(ii) ω° will reveal himself with $m = \omega^{\circ}$ and get v_0 because the maximum following any announcement $m \neq \omega^{\circ}$ is below $v_0 - 2\epsilon$.

(iii) Following $m \neq \omega^{\circ}$, in *Phase Ia* the normal type in indifferent between revealing himself at any point within the current phase, because it will give him the same utility from the start of the current Ia ; delaying revelation until after the current phase is strictly worse.

Step 2 : Player $i > 0$ will not deviate :

(i) during *Phase Ia* : When *Phase Ia* has just been started let player i's utility be v_i^a. Let $T := N - 1 + NP$. The following inequality follows from the recursion structure:

$$\begin{aligned} \frac{v_i^a}{1-\delta} \quad &\geq \quad \lambda_i \left(1 + \delta + \delta^2 + \cdots + \delta^{N-1}\right) + \theta \left\{(-M)\left(\delta^N + \cdots + +\delta^T\right) + \delta^{T+1} \frac{v_i^a}{1-\delta}\right\} \\ &\quad + (1-\theta)\frac{v_i^a}{1-\delta}\delta^N \\ \implies \quad v_i^a \geq &\quad \frac{\lambda_i \left(1-\delta^N\right) - M\theta\delta^N\left(1-\delta^{T-N+1}\right)}{\left\{1-(1-\theta)\delta^N - \theta\delta^{T+1}\right\}} \end{aligned}$$

As $\theta \to 0$, the right-hand side tends to $\frac{\lambda_i\left(1-\delta^N\right)}{\left(1-\delta^N\right)} = \lambda_i$; thus $\exists \theta^*$ s.t $\theta \leq \theta^* \Rightarrow v_i^a \geq \lambda_i - \pi$. The total discounted value at the start of *Phase Ia* comprises the following terms: First we have N periods during each of which which player i expects to get no less than λ_i, given the announced type and the construction of $\rho(\circ)$; with probability θ this is followed by an adjustment phase in which player i gets no less than $-M$ and subsequently play moves to the next Phase Ia, whereas with the remaining probability $1-\theta$ *Phase Ia* is repeated.

(ii) during *Phase Ib* : When *Phase Ib* (the adjustment phase) has just been started let player i's utility be v_i^b. It is easy to see that if i does not have an incentive to deviate at the start of the adjustment phase then she will not do so later, because it is at the start that the player has to be ready to mete out a possibly costly punishment for the greatest number, viz. NP, periods. The minimum utility from carrying this out is $\left\{(-M)\left(1 + \cdots + +\delta^{NP-1}\right) + \delta^{NP}\frac{v_i^a}{1-\delta}\right\}$, where v_i^a is as above. If δ is high enough, this is close to v_i^a. If i deviates, play eventually settles down in an equilibrium that gives η_i to player i. By construction $\eta_i < \lambda_i - \pi$, and $\lambda_i - \pi \leq v_i^a$ from the step above. It therefore follows that a deviation by any $i > 0$ from Phase Ib proves costly.

(iii) Player i does not deviate from III_j or II_j because, in expected value, she ends up losing the reward Δ and any gains from the deviation are wiped out.

(iv) Player i does not deviate from II_i because, in expectation, she is playing her best response anyway; deviation worsens both the current as well as the future payoff. Likewise deviation from III_i is unprofitable because of condition (*). ■

2.5 CONCLUSION

The earlier chapter showed that additional insights were to be gained from studying reputation formation against *multiple patient opponents* ; this chapter shows that the message stands even if two key assumptions are relaxed in turn— that announcements are available and that all non-normal types are *pure* strategy types. The message of both these chapters is that it may not be possible for the patient player to extract the entire surplus while leaving the others with barely their *minmax* values. Let v_0^{min} be the minimum equilibrium payoff of player 0 in the limit when all players are patient and 0 is patient relative to the rest. I characterise an *upper bound* l such that $v_0^{min} \leq l$: Any payoff of the (complete information) repeated game in which 0 gets more than l can be sustained, even when 0 is patient relative to the others and we allow him to commit to mixed(behavioural) strategies of the dynamic game. Furthermore, this bound does not rely on the details of the type-space Ω. To summarise, while a single opponent cannot credibly threaten to punish and thwart a patient player building reputation, the presence of multiple opponents can provide very powerful checks and balances on what reputation is able to guarantee.

APPENDIX

PROOF OF LEMMA 3: By lemma 2 we know that for any $\theta > 0$, $\exists N$ large enough so that for any t and any initial t-period history h^{t-1} we have

$$\Pr\left(\left|\sum_{t=1}^{T}\frac{\xi_t+\xi_{t+1}+...+\xi_{t+N-1}}{N}\right|>\frac{\epsilon}{2}\right)\leq\theta.$$

Now the absolute distance between the disocunted and undiscounted means is, given N,

$$\left|\sum_{t=1}^{T}\frac{\xi_t+\delta_0\xi_{t+1}+...+\delta_0^{N-1}\xi_{t+N-1}}{N}-\sum_{t=1}^{T}\frac{\xi_t+\xi_{t+1}+...+\xi_{t+N-1}}{N}\right|$$

$$\leq\sum_{k=1}^{N-1}\frac{\left(1-\delta_0^k\right)|\xi_{t+k}|}{N}\leq\frac{\left(1-\delta_0^N\right)(N-1)M}{N}<\frac{\epsilon}{2}$$

, when $\delta_0 > \bar{\delta_0}$, where $\bar{\delta_0}$ solves the equation obtained when the last inequality is replaced by an equality; thus the discounted excess payoff over any N periods will with probability $1-\theta$ not exceed $\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon$. ■

Chapter 3

Ideologues beat Idealists: A Paradox of Strategic Voting [1]

> We'd like to vote for the best man, but he's never a candidate. — Frank Hubbard

3.1 INTRODUCTION

It is often the case that voters expect to have more information after the election and before a policy choice is made, as the following hypothetical but plausible scenario shows. A presidential election is in the offing, and the result hinges on one central issue — How best to respond to a country that could pose a threat. Voters could be divided on this issue — some support a direct confrontation (policy-0), while the less hawkish prefer a diplomatic response(policy 1). Voters are aware that their ranking of policies could change after elections are held, but before a policy choice is made; these we call "shocks". These shocks could be weak or inconclusive; voters then react to this "idiosyncratic" shock depending on their personalities. However there is a chance that some very conclusive evidence will come to the fore, either for or against

[1] Co-authored with Vinayak Tripathi.

the said country being a threat, and cause all voters to agree on what the right policy is ; this we call a "common" shock. If it is proved that the enemy was close to developing nuclear weapons, then even the pacifist prefer confrontation; similarly, everybody prefers a diplomatic response when there is no doubt that the adversary is technologically incapable of mounting a serious military threat. Two candidates contest on the following platforms. The first, B, is known to prefer policy 0 come what may; the other, K, is an unbiased candidate who credibly promises to wait until final rankings are formed after the shocks are received, and to then take the action that the majority prefer. Who stands a better chance of being voted to power in a (simple) majority election? One might expect the unbiased candidate to be the natural choice of the population, especially when there is a significant probability of a common shock that makes the committed policy bad for all voters: In the case of a common shock, the unbiased candidate always implements what all voters prefer. We find, perhaps counterintuitively, that voters may prefer the first candidate, thereby committing to a policy rather than waiting to learn their true rankings. This problem is greatly aggravated when voters are rational and their strategies constitute a Nash equilibrium or, equivalently, voters base their decision on the scenario where they are pivotal. This voting equilibrium survives a high chance that new evidence exposes the committed alternative as being undesirable for *all* voters, i.e. a high probability of a common shock away from the policy committed to. Political satirist Frank Hubbard's quip, quoted at the start of the paper, could be turned on its head — (Often) we wouldn't like to vote for the best man even if he were a candidate!

We emphasize that the electorate is faced with a choice between two *candidates*, one of whom always offers a fixed policy from which he derives private benefits, while the other is one who offers a state-contingent plan. Had the electorate's choice had been between two *policies* in the above environment with shocks to rankings, then there would have been a positive probability of each policy being the socially optimal one. What makes our inefficiency quite stark is that the unbiased candidate must, in *any state of the world*, implement the policy that maximises social welfare, but he still loses. When voters choose between two policies each voter chooses the one that he expects to prefer at the next date. As we shall see, this could give rise to inefficient choices as documented by earlier work discussed later in greater depth. A somewhat finer reasoning enters into the decision of the voter when he is strategic and must choose between the two candidates B and K. To evaluate K's plan the voter must now take into account not merely his own ranking of policies but others' rankings and their equilibrium voting strategies; this is because he needs to keep track of where the ex-post social optimum will be. This added element leads the pivotal voter to a new form of inefficiency when he conditions his decision on being pivotal, which fuels his fear of being in a precarious majority even if that is unconditionally very unlikely. This point is elaborated in our examples and proofs.

What are the implications of the common shock for our result ? A common shock captures a situation in which a large number of voters change their rankings in a correlated fashion, as opposed to independently. [2] This feature enriches the insights to be gained from our model: Presence of common uncertainty both adds realism and serves as a test of the robustness of our results. In the event of an common shock, the unbiased candidate provides perfect insurance to all voters in the electorate and therefore a high chance of such an event pushes voters away from the suboptimal or committed candidate. We show that the inefficiency is robust to a high probability of a common shock. When the committed candidate comes to power, the common shock exacerbates the degree of social inefficiency as it creates the possibility of the entire electorate suffering a poor policy choice. On a more reassuring note, in the absence of common shocks, the probability that an elected committed candidate implements

[2]This can be rationalized as follows - each individual's utility function is comprised of a private value component and a common value component. The common shock might significantly alter the common values component of the utility thereby precipitating a correlated shift in rankings.

the socially optimum policy approaches unity for large populations.

The choice between the committed and the unbiased candidate can also be thought of as the choice between picking a policy immediately and waiting until further information makes a more informed choice possible. We show that even with a substantial chance of a common shock, the electorate might choose to act in haste and do the former. Examples of public referenda fit this formulation more naturally. In October 1992, the Swedish Nuclear Fuel and Waste Management Company (SKB), an organization charged with the responsibility of safely disposing nuclear waste, proposed to conduct a study to determine the feasibility of locating a repository. One of the towns that seemed worthy of further investigation was Storuman, in northern Sweden. The proposal polarized the community into those who opposed bringing nuclear waste to Storuman and those who believed that likely economic benefits made the investigation worthwhile. The findings of the SKB would not be binding on the city and if Storuman were deemed feasible it would still be up to the city council to decide, presumably in keeping with public opinion and the interest of the city, whether to allow SKB to actually go ahead with builing a nuclear waste dump. A 1995 referendum asked "whether SKB should be allowed to continue the search for a final repository location in Storuman". The outcome was an overwhelming 'no' (70.5%): the public opted to reject it outright rather than allow more information to be disclosed by a scientific study.

In an article published in the Op-Ed section of the *L.A. Times*[3], Bruce Schulman argues that changing sides has been costly in American politics of late. Candidates spend resources trying to explain away changes in their stand on key issues, from affirmative action to foreign policy. Even fairly incontrovertible evidence of having changed does not dissuade them for arguing otherwise. Schulman suggests that political candidates do not wish to come across as opportunists who pander to the

[3]Schulman, Bruce J. 2007 April. "Beware the politician who won't flip-flop", Op-Ed, L.A. Times.

electorate for political gain. This line of reasoning is also explored by Kartik and McAfee [18] who modify the Hotelling/Downs model of electoral competition to include voters with an explicit preference for candidates with *character* . The results in our paper offer a different explanation for why political candidates might prefer to commit to an ideology rather than update their stands as new information becomes available. We argue that in an environment with changing preferences, the best conceivable flip-flopper, one who adjusts his position to what is best for society at large, cannot expect to win against an ideologue. Office seeking candidates might therefore prefer to be perceived as having ideological biases although the electorate does not intrinsically value this trait.

The results in this paper also offer insights into the type of candidates who *enter* an election. If entering an election is costly and candidates have to choose a plan on which to run, we show that a unbiased candidate may choose not to enter the race. In this case, the electorate will not be presented with the option of voting for the best candidate, as in the tongue-in-cheek quote at the start of the paper.

Our work is closely related to three strands of literature— the first, on status-quo bias against reforms that are put to vote; the second, on pivotal voting; and the third, on spatial competition and the median-voter result. The status-quo bias in reform is well documented — changes that are known to benefit a majority ex-post are not passed ex-ante because some of the would-be winners under the reform vote against the reform; see for example Samuelson and Zeckhauser[16]. Fernandez and Rodrik[4] [FR] provide a rigorous explanation in the context of trade reforms, to our knowledge the first that does not appeal to risk-aversion. Their explanation is based on the identity of the winners within the majority group being unknown at the time of voting. All voters of a group are *ex-ante identical* and hence maximise the expected value of the *group*, behaving in effect like the representative voter of the group. This

[4]* # first drew out attention to this paper, and pointed out a very natural link between our work and theirs.

brings us to the second strand of related literature, which uses the concept of pivotal voting, first introduced in the '*Theory of Voting*' by Farquharson[5]. More recently this difference has been exploited by Austen-Smith and Banks [17] and Feddersen and Pesendorfer [12] to analyse information aggregation in elections. The following comment is intended to avoid legitimate confusion about the role of pivotal voting in our model. Usually the difference between sincere and pivotal voting arises from the fact that the private signal of each voter i affects the valuation of other voters $j \neq i$: When j is pivotal, he can infer the distribution of signals of other voters $i(\neq j)$ and thus updates his rankings conditional on being pivotal. We consider a model with *private values*, i.e. any voter's ranking of alternatives does not depend on others' signals or rankings. But, as we just argued, there is no additional information about voter j's ranking of policies contained in his being pivotal; so why should his vote depend on it at all? The paradox is resolved as follows: When voter j is pivotal, he can infer the probabilities with which the alternative policies will be implemented at the next date by the unbiased candidate K ; this determines voter j 's preferences over the *candidates* and thereby influences the outcome of the voting. Lastly, our work may also be linked to models of spatial voting, notably the pioneering work of Downs and Hotelling. In some senses we provide a framework that argues why the reverse, in a very loose sense to be described later, of the median voter result might hold; in our model a candidate with an extreme position beats an unbiased candidate.

The remainder of our paper is organised as follows: Section 1 puts our contribution in perspective with illustrative examples; Section 2 sets up the model, presents the decision problems of the sincere and pivotal voters, and characterises the resulting equilibria; Section 3 looks at situations where the voting rule at the initial date differs from that used at the subsequent date; Section 4 provides a discussion of the candidate entry problem; and Section 5 concludes.

3.2 ILLUSTRATIVE EXAMPLES

Example 1

We begin with an example that captures the logic of FR. To fix ideas, consider two sectors in an economy — X and Y, with 76 workers in sector in X, and 25 in sector Y. Workers of sectors X and Y are referred to as type-0 and type-1 voters respectively. A simple majority election allows voters to choose whether to stick to the status-quo policy (0) or to implement a reform policy (1). A voter who prefers the policy implemented gets a utility of 1, while others get 0. The shocks are as follows: With probability 0.2 there is a common shock that moves all voters to 1; with probability 0.8 there is an idiosyncratic shock that makes each worker of type-0 become a type-1 with probability p independently of the others. The original model of FR does not have a common shock, but we add it for realism and to facilitate comparison of their model with ours.

Case 1: p is not too small. If $p = .25$, the expected number of people who prefer policy-1 at the next date will be $0.2*101+0.8(25+0.25*76) \approx 56$ which is more than half; so the efficient outcome requires the reform to be passed. But it is not known who the beneficiaries of reform in sector X will be when there is an idiosyncratic shock because all individuals in X are ex-ante identical. Therefore the expected gain to a current type-0 from the reform is $0.2 * 1 + 0.8 * 0.25 * 1 = 0.4$, which is less than the utility from policy-0, $0.8 * 0.75 = 0.6$. The type-0 voter prefers policy 0, and the reform is defeated even though it is welfare improving.

Case 2: p very small, i.e. the idiosyncratic shock is small and does not change the balance of power. Say, $p = .1$. The expected number of people who prefer policy-1 at the next date is $0.2 * 101 + 0.8 * (25 + 0.1 * 76) \approx 46$. Then policy-0 maximises social welfare and all type-0 voters support it because the gain from the reform is $0.2 * 1 + 0.8 * .1 * 1 \approx .28$; there is no inefficiency.

Essentially what each voter does is guard the interests of the group to which he is most likely to belong at the next date. In the above example each type-0 voter is more likely to remain a type-0. Starting with a case where type-0's are in a majority, this is ineffcient if and only if each type-0 is more likely to stay than to switch *and* the type-1's are expected to be in a majority at the next date. Although the model of FR is not directly comparable to ours, their work essentially points out the above source of inefficiency.

Note that there is no way for one voter to be pivotal in the framework above. What if we were to introduce a means by which each voter could be pivotal in the example above? Suppose that instead of there being exactly 76 and 25 workers in the two sectors, we have a model in which a worker was randomly chosen by nature to be of type-0 with probability ≈ 0.75. If type-0's and type-1's vote for different candidates in equilibrium, this assigns a positive probability to each voter being pivotal. Would this change how a typical voter behaves? Note that the voter in FR knows the probability that he will prefer each alternative; he votes for that which generates the highest expected utility. This is a weakly dominant strategy, even when every voter assigns positive probability to being pivotal. Thus the logic outlined in the above example works even if each voter can be pivotal.

As we shall see, our framework allows us to distinguish between pivotal and sincere voting. We show that a combination of two features— voting over candidates instead of policies, and strategic voting — makes the inefficiency more pervasive, although either one alone would not do so. Thus inefficiency may result even when the probability of switching in response to an idiosyncratic shock p is very small, as the following example shows.

Example 2:

Let us now turn to our framework. There are 101 voters, and two alternative policies, 0 and 1. At date-0, the initial date, nature chooses each voter's type, which is the policy he ranks higher. Types are drawn independently — type 0 with probability

.75, and type 1 with probability .25; each voter learns his own type. However there is uncertainty in rankings — voters understand that when it is time to implement a policy at the next date (date 1), their ranking of policies could be different. Each voter gets a utility of 1 if his preferred alternative is implemented, and 0 otherwise. There are two candidates — B, the first one, is known to have private benefits from implementing policy-0; the second, K, is the unbiased candidate who behaves like the social-planner and implements the policy that the majority prefers at date 1.

Next period one of two shocks is possible: With probability 0.2, a common shock causes all voters will switch together towards 1 and with probability 0.8 there is an idiosyncratic shock and each voter switches independently towards 1 with probability $p = .1$ as in case 2 above. We should emphasize here that candidate K is the efficient choice regardless of the shock.

Case a (Sincere Voting): The utility of a type-0 voter from voting B is $1 - 0.2 - 0.8 * 0.1 = 0.72$, since he gets a utility of 1 as long as he stays a type-0. Unconditionally, type-0's are expected to be in the majority at both dates with probabilities close to 1[5]. So if a type-0 voter is sincere and does not condition on being pivotal, his probability of agreeing with K is very close to 1, and he votes for him. There is no inefficiency as in case 2 above.

Case b (Strategic/Pivotal Voting): But the rational voter recognises that his vote matters only when he is pivotal. The utility from voting B is still 0.72, because B chooses a fixed policy. The utility from voting K is now $1 - 0.8 * 0.9 * (1 - (0.9)^{50}) \approx 0.28$. Conditional on the pivotal scenario, the majority will become the minority with probability close to 1 if there is an idiosyncratic shock. The pivotal type-0 voter then ends up in the minority if he is not among those who switch. Conditioning on the knife-edge majority therefore leads him to vote against the reform.

Nature's draw is almost certain to result in type-0's being in a majority at date

[5]It can be checked that if the type-0's are in a majority at date 0, they can become a minority with probability of the order of 10^{-23}.

0; consequently the inefficent choice B is almost certain. A candidate committed to a fixed policy beats the unbiased candidate, even though the probability of idiosyncratic change is very low! If candidate B is elected, the probability that the policy chosen is bad for the everyone is 0.2; this gives a sizeable lower bound on the degree of inefficiency[6]. It might appear at first glance that what the rational voters should try to guard against is the common shock towards 1, when B proves bad for everybody ; but the logic of pivotal voting leads to a surprising conclusion.

In the framework of FR, pivotal and sincere voting yield the same results because voting is over alternative policies not over candidates: Whether or not a voter is strategic, it is a weakly dominant strategy for him to vote the alternative he prefers in expectation. The others' strategies are not relevant given that values are private. The voter is presented with a choice between lotteries whose outcomes are independent of other voters' types and how they vote. As explained in the introduction, our model also has private values, but the strategic element has bite because a pivotal voter can infer the types of other voters and therefore what candidate K is likely to do at the next date.

Inefficiency a la' Fernandez and Rodrik is possible only when the efficient policy hurts the current majority on average; we refer to this as the type I inefficiency. In terms of our model, an idiosycratic swing in rankings must large enough to change the balance of power and to reduce the erstwhile majority to a minority. If p is small then there cannot be any inefficiency. We shall see that, in contrast, the inefficiency displayed by our pivotal voter is more pervasive— it can happen even when idiosyncratic shocks are extremely unlikely to change the balance of power (in expectation). (As noted earlier, the common shocks tilt all voters towards the unbiased candidate.) In terms of our results, our main proposition shows that the inefficiency persists even when the probability p of any voter changing sides is small, and there is a substantial

[6]The majority is most likely to switch only when there is an common shock in favour of policy 1.

probability of a common shock, which helps K. We refer to this as type-II inefficiency.

In summary, our work differs from FR on three counts — model, methodology, and implications. First, in our model voters choose between candidates rather than policy alternatives. Since one candidate offers a policy that is conditional on the electorate's final rankings of policies, a pivotal voter must take into account the voting strategies of others; as a consequence he behaves differently from the sincere voter. Finally, from a substantive point of view the pivotal argument implies that inefficiency remains even when the "idiosyncratic" shock (one that does not affect all voters the same way) is unlikely to precipitate a large change., i.e. p is small.

3.3 THE MODEL

We consider a simple two-period model with voters in $S = \{1, 2, ..., 2n + 1\}$, where $n > 1$, and a set of policies $A = \{0, 1\}$. At date 0, nature draws each voter's type $t_i^0 \in \{0, 1\}$ from a Bernoulli distribution with $\Pr\{t_i^0 = 0\} = q > 1/2$. The voter's type on a particular date specifies the policy he prefers on that date. Elections are held at date 0 as well. There are two candidates to choose from — B and K. Candidate B is known to derive private benefits from the policy 0 [7], while K is an unbiased candidate who promises to maximise social welfare. After the elections, each voter's type changes according to a stochastic process described below. Once each voter's date 1 type t_i^1 is determined, K implements the ex-post social optimum policy if he has been elected at date-0; if B was elected, he chooses policy 0. If $a \in A$ is the policy implemented at date 1, the utility of voter i is given by

$$u_i(a;\ t_i^1) = \begin{cases} 1 \text{ if } a = \ t_i^1 \\ 0 \text{ otherwise} \end{cases} .$$

[7] $q > 0.5$ is without loss of generality as long as there is a candidate who derives private benefits from implementing the policy that is favoured by the majority at date 0.

Fig. 1 below summarizes the temporal structure of the game.

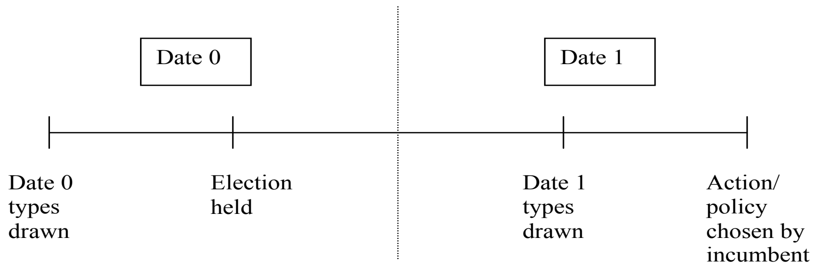

Fig.1: Timeline

Voter's types change over time as new information is revealed to them. Informative signals arrive according to the following process. With probability δ, a public signal favouring one of the policies arrives, forcing everyone to one side; all voters then prefer policy 0 with probability π or policy 1 with probability $1-\pi$. With probability μ, idiosyncratic private signals arrive, leading to independent changes in voters' rankings. These signals favour policy 0 with probability ϕ and policy 1 with probability $1-\phi$. If the signal favours policy 0, each voter of type $t_i^0 = 1$ (henceforth, type-0) switch to policy 0 with probability p independently of the others, while the voters of type-0 stick to their original preferences ; if it supports 1 then all 1 types stay put but each type-0 changes to 1 with probability p independently of the others. Lastly, no information arrives with probability $1-\delta-\mu$; in that case we have $t_i^1 = t_i^0 \forall i$. Later we shall make the simplifying assumption that $\pi = \phi$; this is in no way important for our results and merely permits cleaner algebra and succint interpretations of the derived results. In a realistic case, one would expect δ to be small relative to μ- it is more likely that individuals do not switch en-masse but rather in response to information that each voter chooses to interpret as either strong enough to switch to the other side or too weak to make a difference. All results continue to hold inspite of the relative magnitudes of δ and μ as long as $\delta < \mu$. In the rest of the paper, we refer to the δ-event as being a common shock, and the μ-event as an idiosyncratic shock. Fig. 2 below summarizes the signal structure.

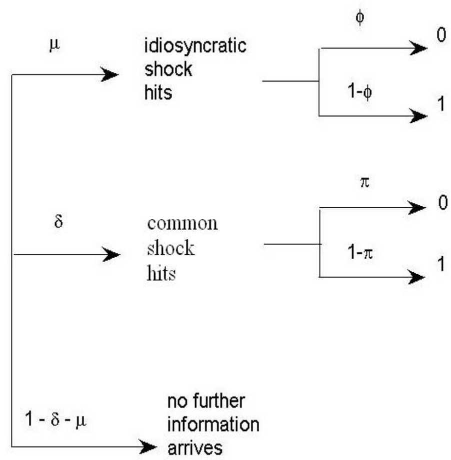

Fig. 2: Schematic Representation of Changing Types

Each voter's type at date 0 is private information; everything else, including the rationality of voters and the stochastic process for the change of types from date 0 to date 1 is common knowledge. It also seems realistic to say that a voter does not know the exact types of the others, but has a general sense of the dispersion in opinion. Since our focus will be on symmetric pure strategy equilibria, assuming types are private information will allow voters to rationally condition on the pivotal scenario.

3.3.1 THE SINCERE VOTER

We use the sincere voter to shed light on the forces that influence the decisions of a pivotal or sophisticated voter. The sincere voter does not condition on the state in which he is pivotal, but instead picks the candidate who, given the unconditional distribution of date 0 and date 1 types, is more likely to agree with him at date 1. He uses a weakly dominated strategy: when he is not pivotal his vote does not affect the result, and when he is indeed pivotal, his vote may not coincide with that of the rational pivotal voter. While falling short of rationality in one of many ways. the sincere voter provides a very useful benchmark against which to compare the results for the pivotal voter, and also facilitates comparison with previous work. An interesting point emerges from the comparison: When all voters are strategic the net outcome could be worse. As we shall see, the sincere voter can only generate one of the two forms of inefficiency discussed below. We should also like to remark at this point that our sincere voter is similar to Harsanyi's rule-utilitarian voter[10]. A rule-utilitarian is one who votes according to the rule that maximises social utility if everyone else follows the same rule. This concept is further extended by Feddersen and Sandroni [6], and by Coate and Conlin [3] to include *group* rule-utilitarians, who choose the action that is best for the *group* when everybody in the group follows it. Our sincere voter chooses like the group-rule utilatarian voter.

Since $q > 0.5$, the vote of the sincere type-0 voter determines the outcome of the election in a large population. We find in Proposition 1 below that the type-0 voter supports K either when (1) type-0's are expected to be in a majority at date 1 , or (2) type-1's are expected to be in a majority at date 1, but the typical type-0 voter is very likely to switch preferences at the next date i.e. p is 'high'. The only case in which he votes B is the one where the majority is likely to be at 1 at the next date but any given voter is very likely to stay put i.e. p is 'low'. In other words, he prefers to commit and safeguard his interests today as he might not have enough support to do so at the next date.

PROPOSITION: *When $q \in (0.5, 1)$ and n is large enough, sincere type-0 voters vote B if and only if $q(1-p) < 1/2$ and $p < \frac{1}{2}(1 - \frac{\delta}{\mu})$. Otherwise they vote K. The sincere type-1 voters always vote K. When type-0 voters support B, the probability that B wins goes to 1 as $n \to \infty$.*

PROOF: The expected utility of a sincere type-0 voter i when B is elected is given by

$$U_i(B,0) = 1 - \mu(1-\phi)p - \delta(1-\pi)$$

; the second term corresponds to the loss incurred when the voter sways to an idiosyncratic 1-signal and the third term is the loss due to an common 1-signal. When

K is elected, the expected utility is

$$U_i(K,0) = 1 - \mu(1-\phi)p\Lambda_{01} - \mu(1-\phi)(1-p)\Lambda_{10} - \mu\phi\Pi_1$$

, where Λ_{01} is the probability that the date 1 majority is at 0 when type-0 voter i switches to 1 in response to an idiosyncratic 1-shock, Λ_{10} is the probability that the date 1 majority is at 1 when voter i ignores the idiosyncratic 1-signal and Π_1 is the probability that the majority stays at 1 despite an idiosyncratic 0-signal.[8] The three negative terms correspond to the three potential sources of loss under K. Note that losses can only be incurred under K when idiosyncratic signals arrive. Conditioning on an idiosyncratic 1-signal, the probability that an arbitrary voter supports policy 0 at date 1 is $q(1-p)$ for large n. If we define the random variable $X_{01} \sim \text{Binomial}(2n, q(1-p))$ as the number of voters (barring one) who support policy 0 at date 1 following an idiosyncratic 1-signal, then $\Lambda_{01} = \Pr\{X_{01} > n + 1\} = \Pr\{\frac{1}{2n}X_{01} > \frac{1}{2} + \frac{1}{2n}\}$. The Weak Law of Large Numbers guarantees that $\lim_{n\to\infty} \Pr\left\{\left|\frac{1}{2n}X_{01} - q(1-p)\right| < \epsilon\right\} = 1$ for any $\epsilon > 0$. If $q(1-p) > 0.5$, there exists an n large enough and an $\epsilon > 0$ so that

$$\Pr\left\{\left|\frac{1}{2n}X_{01} - q(1-p)\right| < \epsilon\right\} \leq \Pr\left\{\frac{1}{2n}X_{01} > \frac{1}{2} + \frac{1}{2n}\right\} \equiv \Lambda_{01} \leq 1.$$

It follows that

$$q(1-p) > 0.5 \Rightarrow \lim_{n\to\infty} \Lambda_{01} = 1 \text{ and } \lim_{n\to\infty} \Lambda_{10} = 0 \text{ ; } q > 0.5 \Rightarrow \lim_{n\to\infty} \Pi_1 = 0$$

For large n and $q(1-p) > 0.5$, the type-0 voter then votes K, the unbiased candidate for all $\delta \geq 0$. When $q(1-p) < 0.5$, the date 1 majority is expected to be at 1 and for large n, $\Lambda_{01} \to 0$, $\Lambda_{10} \to 1$ and $\Pi_1 \to 0$. The type-0 voter then votes K if

[8] $\Lambda_{01} = \sum_{j=n+1}^{2n} \binom{2n}{j}\theta^j(1-\theta)^{2n-j}, \Lambda_{10} = \sum_{j=0}^{n-1} \binom{2n}{j}\theta^j(1-\theta)^{2n-j}, \Pi_1 = \sum_{j=0}^{n-1} \binom{2n}{j}\psi^j(1-\psi)^{2n-j}, \theta = q(1-p), \psi = q + (1-q)p$

$p > \frac{1}{2}\left(1 - \frac{\delta(1-\pi)}{\mu(1-\phi)}\right)$. Under $\pi = \phi$, this reduces to $p > \frac{1}{2}\left(1 - \frac{\delta}{\mu}\right)$.

Let us now turn to the sincere 1-voter. His expected utility from B is given by

$$U_i(B,1) = (\delta\pi + \mu\phi p)$$

; he gets a utility of 1 iff he switches to 0 himself, in response to either an idiosyncratic or an common 0-signal. His utility from K is

$$U_i(K,1) = \delta + \mu\phi p\Pi_{00} + \mu\phi(1-p)\Pi_{11} + \mu(1-\phi)\Lambda_1$$

, where Π_{ab} is the probability that, following an idiosyncratic 0-signal, the majority is at a when our voter is at b; $\Lambda_1$is the probability that, following an idiosyncratic 1-signal, the majority is at 1. By arguments similar to the ones made for the sincere 0-voter above,

$$q > 0.5 \Rightarrow \lim_{n\to\infty} \Pi_{00} = 1 \text{ and } \lim_{n\to\infty} \Pi_{11} = 0$$

$$q(1-p) > 0.5 \Rightarrow \lim_{n\to\infty} \Lambda_1 = 0 \text{ ; } q(1-p) < 0.5 \Rightarrow \lim_{n\to\infty} \Lambda_1 = 1.$$

When $q(1-p) < 0.5$, we can therefore reduce the decision for large n to one of comparing $\delta\pi + \mu\phi p$ with the quantity $(\delta + \mu(1-\phi) + \mu\phi p)$ that he gets from K. Since $\delta(1-\pi) + \mu(1-\phi) > 0$, he votes K. When $q(1-p) < 0.5$, he compares $\delta\pi + \mu\phi p$ from B with $\delta + \mu\phi p$ from K and votes K. Thus the sincere type-1 voter votes K for all values of $q \in (0.5, 1)$and all $p \in (0, 1)$.

When either (1) $q(1-p) > 0.5$, or (2) when $q(1-p) < 0.5$and $p > \frac{1}{2}\left(1 - \frac{\delta}{\mu}\right)$, both the type-0's and 1's vote K and he wins with probability 1. When $q(1-p) < 0.5$ but $p < \frac{1}{2}\left(1 - \frac{\delta}{\mu}\right)$, the type-0's vote B while the 1's vote K; since $q > 0.5$, B is the more likely winner; his exact probability of winning is $\Sigma_{k=n+1}^{2n+1}\binom{2n}{k}q^k(1-q)^{2n-k}$, which is greater than 0.5 for all n and tends to 1 as $n \to \infty$. ■

Fig. 3 below summarizes the behaviour of the type-0 voter for large enough n. In regions II and IV, the type-0 voter elects K as he expects to remain in the majority if there is an idiosyncratic shock, and has nothing to lose by voting K; when there is an aggregative 1-signal, he is better off with K as B would still continue to implement the alternative 0, which gives B private benefits; with an common 0-signal, both B and K implement 0. In region I, the date 1 majority is expected to prefer policy 1, but since p is high each type-0 expects to switch and be in the subsequent majority. So he votes for K, who always picks the right alternative when there is an aggregative shock. Finally in region III, the sincere type-0 voter picks the socially suboptimal candidate B because the majority is likely to prefer 1 at the next date, but given that p is small he would probably stay put at 0.

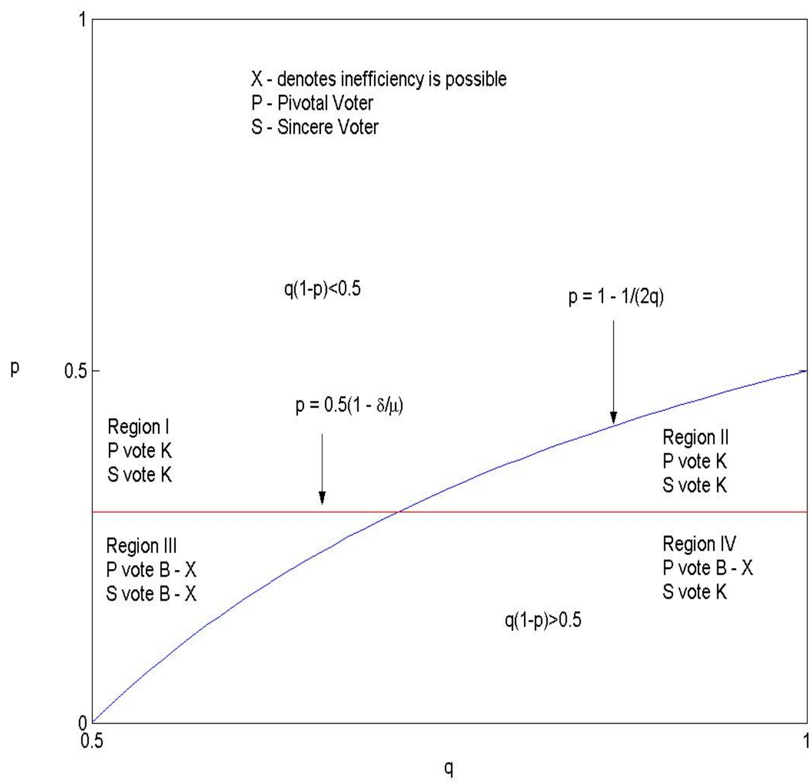

Fig. 3: Voting Behaviour of the Type-0 Voter

The sincere voter is nothing but the representative agent of the group, one who maximises the value of the group[9]This is easy to see - all agents of a group are ex-ante identical, and the probabilities that appear in the decision of the sincere voter are the expected proportions in the decision of the representative agent. The probability of switching, for example, is now to be interpreted as the expected proportion of type-0's who switch when there is an idiosyncratic 1-signal.

3.3.2 THE PIVOTAL VOTER

The situation is very different when we require the strategies to constitute a Nash equilibrium. We conjecture the following Nash equilibrium in pure strategies : each voter of type-0 votes B, and all type-1's vote K; then we solve for the range of parametric values where this is indeed the case.[10] Let $U_i(c \mid c_0, c_1;\ t_i^0,\ \text{piv})$ denote the utility of the pivotal voter of date 0 type t_i^0 when he votes for candidate c, the type-0's vote for candidate c_0 and the type-1's vote for candidate c_1. Consider a pivotal type-0 voter. When he is pivotal the utility of voting for B is the same as that for the sincere voter:

$$U_i(B \mid B, K;\ 0,\ \text{piv}) = 1 - \mu(1-\phi)p - \delta(1-\pi)$$

The utility of voting for K is now different :

$$U_i(K \mid B, K;\ 0,\ \text{piv}) = \delta + (1-\mu-\delta) + \mu\phi + \mu(1-\phi)\left\{(1-p)^{n+1} + p\right\}$$

Utility from K is 1 whenever there is an common signal, no signal, or an idiosyncratic 0-signal. If an idiosyncratic 1-signal arrives, the type-0 voter gets 1 if nobody switches or if he himself switches. The pivotal voter prefers B when

[9]The sincere voter's behaviour is akin to the notion of group rule utilitarianism introduced by Harsanyi.

[10]We ignore trivial equilibria in which all voters support the same candidate. Our focus is on symmetric equilibria in which all voters of a type vote the same way.

$$\begin{aligned} U_i(B \mid B, K;\ 0,\ \text{piv}) &> U_i(K \mid B, K;\ 0,\ \text{piv}) \\ \Leftrightarrow \delta(1-\pi) + \mu(1-\phi)p &< \mu(1-\phi)(1-p)\{1-(1-p)^n\} \end{aligned} \tag{3.1}$$

Under the assumption $\pi = \phi$, the condition reduces to

$$f(p, n) = 1 - 2p - (1-p)^{n+1} > \frac{\delta}{\mu}.$$

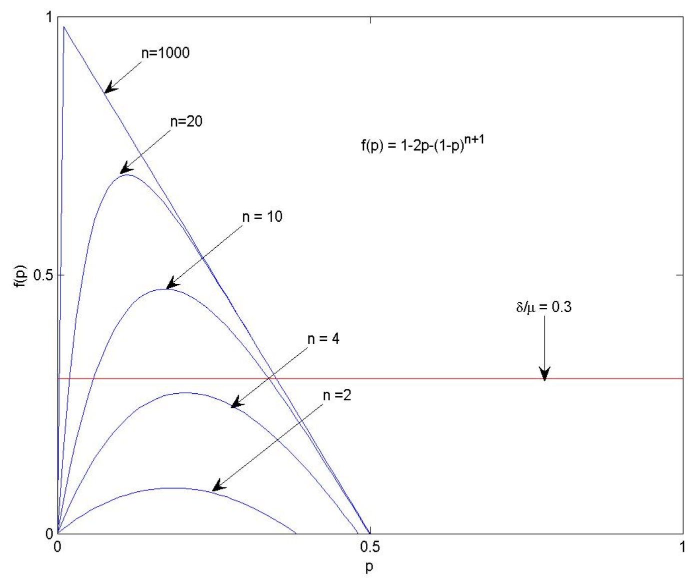

Fig 4: Simulation Results for "Small" n

Fig. 4 above shows how $f(p, n)$ compares to a threshold of δ/μ for different values of n. Since $f(n+1, p) > f(n, p)$ for all p in $(0, 1)$,the range of values of p for which the committed candidate wins is growing with n; as $n \uparrow \infty$, the negative term $(1-p)^{n+1}$ goes rapidly to 0, and the condition reduces to $1 - 2p > \frac{\delta}{\mu} \Leftrightarrow p < \frac{1}{2}(1 - \frac{\delta}{\mu})$.[11] There

[11] If we turn to the case when there is no common shock, i.e. $\delta = 0$, then 0's vote B if $(1-p)^{n+1} < 1 - 2p$. Using a quadratic Taylor series expansion, a sufficient condition for this is $1 - (n+1)p + \frac{(n+1)n}{2}p^2 < 1 - 2p$, or $\frac{(n+1)n}{2}p < n - 1$. i.e. $p < \frac{2(n-1)}{n(n+1)}$. For 7 voters, for example, this effect is observed for $p < 1/3$. Thus with $\delta = 0$, the bias towards commitment is very much a reality even with relatively few voters.

exists a $p > 0$ satisfying the above if $\delta < \mu$, i.e. if the probability of an common switching signal is less than that of a idiosyncratic signal. The only requirement is that a signal that causes the entire population to switch to one side is less likely than a signal that causes voters to switch idiosyncratically, - surely a reasonable assumption. So δ may be quite large, even close to but less than one-half, provided that it is less than μ.

For n large, (**??**) reduces to

$$\delta(1-\pi) + \mu(1-\phi)p < \mu(1-\phi)(1-p) \tag{3.2}$$

The LHS of (3.2) is the loss from voting for B and getting him elected - the first term is the loss when the entire population switches to 1 at date 1, the second is when voter i responds to an idiosyncratic signal and finds himself on the wrong side vis-a-vis B. The RHS is the loss when K is elected. Note that the pivotal voter reacts very differently to the possibility of an idiosyncratic switch depending on who is in power- B or K. When B is in power, the loss is when voter i himself switches whether or not others switch. With K in power, i no longer fears switching even if others don't; what he fears is staying put when others switch.

PROPOSITION 2: *When $q \in (0.5, 1)$ and n is large enough, there exists a Nash equilibrium with pivotal voters in which all type-0's vote B and all type-1's vote K when $p < \frac{1}{2}(1-\frac{\delta}{\mu})$. B wins with a probability that tends to 1 in large populations. When $p > \frac{1}{2}(1-\frac{\delta}{\mu})$ all voters (type-0 and 1) prefer K to B.*

PROOF: The argument above shows that for large enough n, the pivotal type-0 voters prefer B if $p < \frac{1}{2}(1-\frac{\delta}{\mu})$ and the type-1 voters conform to the conjectured equilibrium strategy. Since $q > 0.5$, the Weak Law of Large Numbers ensures that almost surely, a majority of voters support 0 on election day and therefore B wins.

Finally to show that the above conjecture indeed gives us a NE, we show that the pivotal type-1's will vote K. Expected utility from B is

$$U_i(B \mid B, K;\ 1,\ \text{piv}) = \delta\pi + \mu\phi p$$

, while that from K is

$$U_i(K \mid B, K;\ 1,\ \text{piv}) = (1-\mu-\delta) + \delta + \mu(1-\phi) + \mu\phi(1-p)^{n+1} + \mu\phi p.$$

Under $\mu = \phi$, type-1 's vote K if $\phi(\mu + \delta) < 1$, which necessarily holds. ■

For large n, there are two forms of inefficiencies illustrated above. The first, which corresponds to region III of Fig. 3 and is exhibited by both sincere and pivotal voters, was discussed in Section 2.1. The difference between the sincere and the pivotal voter is in region IV: the pivotal voter prefers the committed candidate even when, following an idiosyncratic shock, he expects to remain in the majority. What drives the fear of the pivotal voter is that his decision is conditional on himself being pivotal, thereby unravelling the effect of q. In contrast to the decision of the sincere voter for large n, his decision depends only on the value of p and not that of q. The interaction among pivotal voters enters through the size of the population: when n is large it is almost certainly the case that, starting from a pivotal situation, the pivotal 0-type voter will be in minority if he does not switch following an idiosyncratic signal. The proposition below summarises this.

PROPOSITION 3 *When* $p < \frac{1}{2}(1 - \frac{\delta}{\mu})$ *and* $p < 1 - \frac{1}{2q}$ *, sincere voting results in election of the unbiased candidate, while pivotal voting almost surely results in the election of the committed candidate. If B wins, he implements a suboptimum policy with a probability that is bounded below by the probability of an common 1-type shock* $\delta(1-\pi) > 0$. *All sincere voters vote K, whereas the pivotal type-0 and 1 vote B and K respectively. Thus for large* n, *B wins with a probability arbitrarily close to unity if voters are fully rational; K wins if they are sincere.*

REMARK *For all n, not necessarily large, and $q > 0.5$ the probability of an inefficient decision is bounded below by $\delta(1-\pi)/2$. This follows since B wins with a probability bounded below by* 0.5 *for all n.*

Finally, returning to Fig. 4 , we note that for small n, it is not the case that the pivotal type-0's prefer B for all values of p less than a certain threshold. For very small p the condition reduces to $\delta < 0$, which is impossible. What is the reason for this difference? Recall that the pivotal agent's fear is of being left behind - of not switching to the other side while at least one other person on his side defects and destroys a fragile majority. But the fewer the voters the less likely is this fear, and so the dominant fear is that of an common 1-shock that would render B undesirable for all voters.

3.4 VARYING THE VOTING RULE

The previous sections proceeded under the assumption that the voting rule used at date 0 is the same as the decision rule used at date 1 by the unbiased candidate K. Recall that an interpretation of our framework, one that we mention earlier, is the choice between acting now or waiting; with this interpretation it is indeed natural to suppose that the voting rule at 0 and K 's rule at 1 are the same. But if we think of it as an electoral contest, one is naturally led to investigate the properties when the two rules are different. This section accordingly looks at an $m-rule$ at date 0, to be defined shortly. In this section, we examine the behaviour of the pivotal voter for a range of such voting rules.

As one would expect, increasing m from $\frac{1}{2}$ makes it difficult for B to win, and helps mitigate the inefficiency generated by the pivotal voter. However the inefficient equilibrium turns out to be robust in a large range of (m,p) values. To see this, let us re-examine the conjectured equilibrium in which type-0's vote B and type-1's vote

K, and analyze the pivotal type-0 voter's decision. Recall that $q > 0.5$ means that in the symmetric equilibrium the vote of the type-0 voters is the deciding vote. Under the new voting rule, the pivotal voter conditions on the state of the world in which k voters are of type-0 , where $\frac{k}{2n+1} < m \leq \frac{k+1}{2n+1}$. If $m(1-p) > \frac{1}{2}$ and n is large, by the law of large numbers we know that it is highly likely the majority at date 1 will remain at 0. The pivotal voter is therefore not conditioning on a precarious majority and this allays his fear of being left behind while the majority switches to policy 1 at the next date. Under this condition we thus find that the inefficient equilibrium conjectured above fails to exist. By providing a buffer between the point the pivotal voter conditions upon and the simple majority, the $m-rule$ reduces the type-0 voter's incentive to protect himself by voting for the committed candidate B. When the above condition is not met the inefficient symmetric equilibrium survives.

More formally, the utility of the pivotal voter of type-0 when he votes B , given that all 0's vote B and K 's vote 1 following the dictates of the equilibrium, is identical to that in the previous section and is given by

$$U_i\left(B|B,K;0,piv\right) = 1 - \mu\left(1-\phi\right)p - \delta\left(1-\phi\right)$$

, where we have assumed $\phi = \pi$ for simplicity. The utility of the pivotal voter above from voting K can be written as

$$U_i\left(K|B,K;0,piv\right) = \delta + \left(1-\mu-\delta\right) + \mu\phi + \mu\left(1-\phi\right)\left\{p\Theta_1 + (1-p)\Theta_0\right\}$$

, where Θ_a is the probability that the majority prefers policy $a \in \{0,1\}$ at date 1, conditional upon starting at date 0 from a situation in which $k+1$ voters are of type-0 and the rest of type-1. The four terms correspond to the situations in which the pivotal voter i of type $t_i^0 = 0$ gets a utility of 1 by voting B and thereby electing him. The first term δ is for an common shock, when all voters agree on the policy and B

implements it ; the next is when no additional information arrives and i continues to be in a $m-$supermajority and thus in a simple majority at date 1; the third term is for the common shock towards 0 , when the majority for 0 is bolstered; the last terms is for the indiosyncratic shock towards 1— i gets 1 iff he switches and the majority swings to 1 or if he stays put and so does the majority. Depending on whether p is large or small relative to the value m , Θ_0 or $\Theta_1$is much larger than the other in the limit. Let us first consider the case when $p < 1 - \frac{1}{2m}$. By the Weak Law of Large Numbers, for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\text{fraction of popln. supporting policy 0 at next date } - m(1-p)| < \epsilon) = 1$$

It follows then that if we choose a small enough ϵ then $\lim_{n\to\infty}\Theta_0 = 1$ and $\lim_{n\to\infty}\Theta_1 = 0$. Therefore, the utility from voting for K converges to

$$\lim_{n\to\infty} U_i\left(K|B,K;0,\,piv\right) = 1 - \mu\left(1-\phi\right)p$$

For the pivotal voter to prefer candidate B, it is therefore necessary for $\delta\left(1-\phi\right) < 0$. Since this is not true, the conjectured equilibrium does not exist when the electorate is large and $p < 1 - \frac{1}{2m}$.

When $p > 1 - \frac{1}{2m}$, a similar argument gives $\lim_{n\to\infty}\Pi_0 = 0$ and $\lim_{n\to\infty}\Pi_1 = 1$. The utility from voting for K then converges to

$$\lim_{n\to\infty} U_i\left(K|B,K;0,\,piv\right) = 1 - \mu\left(1-\phi\right)\left(1-p\right)$$

In this case, the pivotal voter prefers voting B when $\delta < \mu(1-2p)$, or equivalently $p < \frac{1}{2}(1-\frac{\delta}{\mu})$. Note that this constraint is identical to the one derived for the simple majority rule. Fig. 5 shows how the pivotal $type-0$ voter's relative preference for each candidate varies with the parameter p for different $m-rules$. For each m, the voter prefers candidate B to K for values of p where the curve lies above 0. Note that as $m \to \frac{1}{2}$, the inefficiency is possible for smaller and smaller values of p. When $m = \frac{1}{2}$, any $p > 0$ can give rise to the inefficient equilibrium for large enough electorates; this is the content of the previous section.

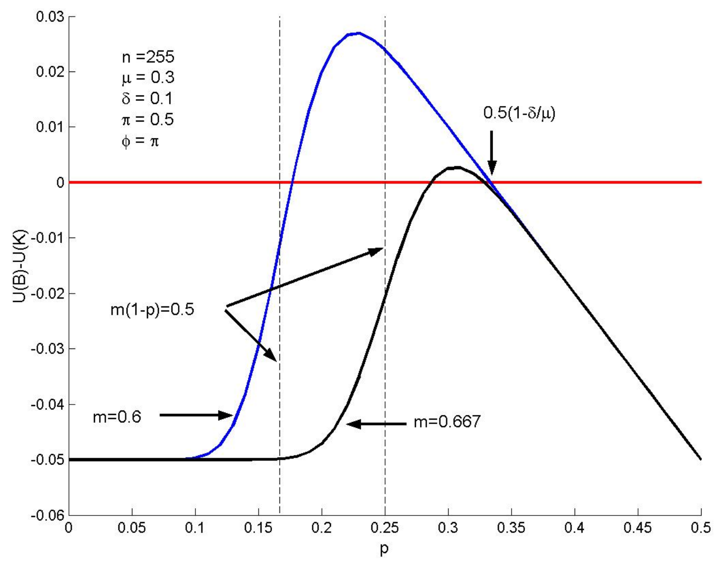

Fig. 5: Simulation Results for m-Rules with $m > \frac{1}{2}$

PROPOSITION 4: *The inefficient equilibrium exists for large electorates when* $1-\frac{1}{2m} < p < \frac{1}{2}(1-\frac{\delta}{\mu})$. *The set of values of* p *which supports the inefficient equilibrium shrinks as* m *increases. If* $m > \frac{\mu}{\mu+\delta}$, *the inefficient equilibrium does not exist for any value of* p.

PROOF: We have already verified that the pivotal $type-0$ voters will conform to the behaviour of the inefficient equilibrium when p is in the range specified. Now we show that the pivotal $type-1'$s will vote K. Expected utility from B is

$$U_i(B \mid B, K;\ 1,\ \text{piv}) = \delta\pi + \mu\phi p$$

, while that from K is

$$U_i(K \mid B, K;\ 1,\ \text{piv}) = \delta + \mu(1-\phi)\Pi_1 + \mu\phi p$$

Under $\phi = \pi$, and $p > 1 - \frac{1}{2m}$, $type - 1$'s vote K if $\phi < 1$, which necessarily holds.■

It follows immediately from the last inequality in the proposition above that the inefficient equilibrium cannot arise for the unanimity rule ($m = 1$). While our framework is not directly comparable to the information aggregation models, it might be interesting to note that this result contrasts with the inferiority of the unanimity rule documented previously.

3.5 IMPLICATIONS FOR CANDIDATE ENTRY : A DISCUSSION

So far our work proceeded under the assumption that candidate preferences were common knowledge. If we now choose to move back one step in time, we might ask which candidates actually enter an election. Let $c > 0$ denote the cost of entering an election; this includes the cost of filing nomination, campaigning, etc. The candidate could be either one who cares about his legacy and when in office picks the alternative that the majority prefer, or a partisan who benefits from one of the two alternatives independently of the electorate's rankings. Let us consider a simple extension of the model where q has not been revealed when the candidates decide to enter. Suppose q is equally likely to be 0.25 and 0.75 and let b denote the benefits from being in office for any type of candidate. If there is a potential candidate of each type and $b > 2c$, then in equilibrium one 0-candidate and a 1-candidate will both choose to enter. When $q = 0.25$ the 1-candidate wins with a probability arbitrarily close to 1 for a large enough n ; with $q = 0.75$ the 0-candidate wins. The legacy candidate will therefore choose not to enter the fray as his expected gain from entry will fall short of the cost $c > 0$ — He cannot win the election irrespective of what value of q is drawn.

As political satirist Frank Hubbard once said, “We’d all like to vote for the best man but he’s never a candidate”. We see that extreme candidates do better than one who is completely unbiased and proposes to implement the ex-post majority’s preferred policy. In a loose sense this hints that gradual resolution of uncertainty might have something to do with the entry of biased candidates.

3.6 CONCLUSION

Many papers look at information aggregation in voting and the role of pivotal voting. They almost always have the feature that voters’ rankings of policies do not change betweent the time they vote and the time a policy is implemented. The role of the electoral process is to aggregate their private signals. We relax this framework and allow rankings to change; this leaves the electoral system prone to widespread inefficiency. This paper illustrates two forms of inefficiency. When voters who are in a majority today are more likely to be in a minority tomorrow, they oppose social-welfare improving policies. This requires a probability of idiosyncratic switching large enough to reduce the ex-ante majority to an ex-post minority. Perhaps a large range of electoral situations is better described by a model in which the probability of voters changing idiosyncratically is small. This, one might even assert, is the rule rather than the exception. We should hope that in such a case the inefficiency will be mitigated, if not eliminated. We argued above that this is not the case — In the unique (informative) symmetric Nash equilibrium, voters prefer to elect the ideologue rather than elect an unbiased candidate, who waits for all information to be revealed and thereafter takes the optimal decision. The key to understanding this paradoxical result is that the *pivotal* voter finds himself in a fragile majority that is easily overturned; even though such a situation is (unconditionally) unlikely, he bases his vote on this situation and commits to the alternative that he currently prefers.

This continues to hold even if there is a large chance that everybody will dislike the committed candidate's choice due to an common shock.

Bibliography

[1] AUSTEN-SMITH , DAVID, AND JEFFREY BANKS. 1996 “Information Aggregation, Rationality and the Condorcet Jury Theorem,” *American Political Science Review*, 90(1): 34-45.

[2] CALLANDER. 2008. ”Political motivations,” *Review of Economic Studies.*

[3] COATE, STEPHEN, AND MICHAEL CONLIN. 2004. “A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence” *American Economic Review,* 94(5): 1476-1504.

[4] FARQUHARSON, ROBIN 1969. *Theory of Voting*. New Haven: Yale University Press.

[5] FEDDERSEN, TIMOTHY, AND WOLFGANG PESENDORFER. 1997. “Voting Behaviour and Information Aggregation in Elections with Private Information” *Econometrica*, *65*(5): 1029-1058.

[6] FEDDERSEN, TIMOTHY, AND ALVARO SANDRONI. 2006. “A Theory of Participation in Elections” *American Economic Review*, 96(4): 1271-1282.

[7] [FR] FERNANDEZ, RAQUEL, AND DANI RODRIK. 1991. “Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty” *American Economic Review, 81*(5): 1146-1155.

[8] Friedenberg, Amanda and Ethan Bueno de Mesquita 2008. "Ideologues or Pragmatists?" *mimeo*

[9] Hansson, Ingemar, and Charles Stuart. 1984. "Voting Competitions with Interested Politicians: Platforms do not Converge to the Preferences of the Median Voter," *Public Choice*, 44(3): 431-441.

[10] Harsanyi, John. 1980. "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior," *Theory and Decision*, *12*(2): 115-133.

[11] Sumon Majumdar; Sharun W. Mukand 2004. "Policy Gambles *The American Economic Review*, Vol. 94, No. 4.

[12] Margolis. 2008. "Pivotal Voting and the Emperor's New Clothes".

[13] Kartik, Navin, and R. Preston McAfee. 2007. "shocking Character in Electoral Competition", *American Economic Review*, 97(3): 852-870.

[14] Myerson, Roger. 2000 "Large Poisson Games", *Journal of Economic Theory*

[15] Razin, Ronny 2004. "Signalling and Election Motivations in a Voting Model with Common Values and Responsive Candidates" Econometrica. 71(4).

[16] Samuelson, William, and Richard Zeckhauser. 1988. "Status quo bias in decision making*". Journal of Risk and Uncertainty,* 1(1): 7-59.

[17] Stokes, Donald E. 1963. "Spatial models of party competition", *The American Political Science Review.*