

# FREE RIDING AND INCENTIVES TO INVEST IN THE REPUTATION OF AN ANONYMOUS GROUP

Dragan Filipovich\*

Centro de Estudios Económicos, El Colegio de México, A. C.

(Received 12 July 2001, accepted 14 January 2002)

---

---

## Abstract

This paper studies incentives to invest in “group reputation” when some agents are unable to recognize their partners from past interactions (*i.e.*, anonymity). It does this by embedding the Kreps-Wilson (1982a) model of reputation and entry deterrence in a random-matching game (instead of just looking at isolated bilateral interactions). Examples are presented of how this affects investments in reputation under differing assumptions on how information is shared amongs agents.

## Resumen

Este trabajo estudia los incentivos para invertir en la reputación de un grupo cuando algunos de los agentes no reconocen a sus socios anteriores (*i.e.*, existe anonimato). Para lograr este propósito, se utiliza el modelo de Kreps y Wilson (1982a) sobre reputación y restricciones de entrada en un juego de asociación aleatoria (en lugar de considerar en forma aislada interacciones bilaterales). Por último, se desarrollan varios ejemplos para mostrar cómo se modifica la inversión en reputación cuando la información disponible se comparte de distintas formas entre los agentes.

*JEL Clasification: C72, D82, and D83.*

*Keywords: Group Reputations, Anonymity, and Random Matching.*

---

\* Centro de Estudios Económicos, El Colegio de México A. C., Camino al Ajusco 20, 10740 México, D. F., Telephone (52)54493050, E-mail: dfilipovich@colmex.mx

The author is very grateful to two anonymous referees for their comments.

## 1. Introduction

Reputation is often shared among individuals, *i.e.*, the actions of any individual in a given group influence not only how people interact with that individual, but also how they interact with the remaining members of the group. For example, when a customer gets less than satisfactory service from a company's employee, he or she reacts by switching companies, rather than by trying to avoid dealing with the offending employee (not a practical option, most of the time). In fact, it seems only a slight exaggeration to claim that practically any easily identifiable group (ethnic, political, professional, *etc.*) "shares a reputation" in this general sense.

Of course, what causes a group to "share" a reputation varies considerably. In some cases, the fact that a member of the group behaves in a certain way might make it more likely that other members of the group will behave in the same way (*e.g.*, bad service resulting from lax supervision). In other cases, people might quite deliberately retaliate against the collective, so as to induce it to discipline its individual members (despite being able to identify individual offenders).

In yet other cases, it might simply be that outsiders cannot identify individual members within the group. They avoid dealing with the group in order to avoid dealing with the individuals who provided inadequate service. For example, a person might remember that the driver was mildly alcoholized last time she took a cab at a particular corner-stand, yet might have forgotten who the driver was. She might quite reasonably opt not to patronize the corner-stand at all, just to be on the safe side.

This paper presents a model of the last class of situations, specifically, a model in which agents are unable to recognize the group members they have interacted with in the past ("anonymity"), and in which, as a consequence, beliefs about ex-ante unobservable characteristics (*e.g.*, propensity to have a few drinks during working hours) attach to the whole group, rather than to individuals. The key issue is whether the members of the group will have incentives to invest in the group's reputation (which effectively becomes a sort of public good), *i.e.*, to engage in costly activities in the short run in order to influence other agents' beliefs about the unobservable characteristics of the group in the longer run. Or putting it in another way, we ask to what extent, if at all, does the presence of such "reputational externalities" lead agents to free ride on each other, and to what extent does eventual free-riding interfere with the formation of group reputations of this kind?

To study the above questions, we embed a repeated entry-deterrence stage-game with imperfect information à la Kreps-Wilson (1982a) in an anonymous random matching game (two long-lived incumbents matched randomly each period with two one-period lived entrants who share their experiences with their successors). The notion of reputation we work on is Kreps and Wilson's: entrants have beliefs about the incumbent's type (compulsive fighter or potential compromiser), which they update according to their predecessors' experience.

Incumbents who are potential compromisers have incentives to fight (even when this is not the optimal short-run action) in order to convince entrants that they are compulsive fighters, and, thus, deter future entry. The additional twist is that now entrants' successors are not able to tell apart the incumbent who fought from the incumbent who accommodated, and so will tend to "average" their beliefs. This, in turn, will reduce incumbents incentives to sacrifice current remunerations for future ones (*i.e.*, to deter entry). If moreover, contemporaneous entrants share information (*i.e.*, both know everything that has occurred in the past, rather than only the experiences of their "family"), incumbents might free ride on each others' investments in reputation.

The first part of this paper presents an scenario with information-sharing among contemporaneous entrants (but no information-sharing amongst incumbents, this means that incumbents do not know each other's types), in which for certain parameter constellations (under which there would be reputation building in Kreps and Wilson), there will not be any here, and this regardless of the horizon of play (in contrast to Kreps and Wilson's classic conclusion that there will always be reputation building given a sufficiently long horizon). As will become apparent, this is a direct consequence of there being "free-riding" due to information-sharing.

The second part of this paper presents a two-period analysis of what happens when there is information-sharing both by entrants and incumbents (*i.e.*, incumbents now know each other's types). Interestingly, the two-period equilibrium here differs only marginally from the corresponding equilibrium in the Kreps-Wilson model for a whole range of parameter values, suggesting that simultaneous information-sharing by entrants and incumbents might generally help neutralize negative free-riding effects on reputation. By the way, it emphasizes the importance of the exact pattern of information-sharing for the formation of reputations in collective environments.

Finally, one should point out that "anonymity" is a common assumption in both the theoretical and experimental literature. It is often made in random matching models to preclude reputation-building, and more generally, any kind of repeated-game effects. The results in this paper constitute further evidence that this might not always work (see Ellison, 1994). A discussion of the relevant literature and an outline of the remaining sections complete this introductory section.

## 2. Literature Discussion

This work draws mainly on the seminal contributions of Kreps and Wilson (1982a), and Milgrom and Roberts (1982a)<sup>1</sup>. The model used here is actually an extension of the Kreps-Wilson (1982a) model of entry deterrence through reputation. In the present extension, two (one-period lived) entrants are matched

---

<sup>1</sup> For an overview of the literature originating in these contributions, see Fudenberg and Tirole (1992).

each period one-on-one and randomly with two (long-lived) incumbents, instead of one short-lived entrant being matched with the same long-lived incumbent each period, as in Kreps-Wilson. Another difference is that, here, tough incumbents are robots rather than just agents who like to fight. Despite these differences, incentives to invest in reputations arise here in pretty much the same way as in Kreps-Wilson<sup>2</sup>.

A few words about the choice of approach. Much more general versions of the results in the Kreps-Wilson paper just mentioned have been obtained by Fudenberg and Levine (1989, 1992) using an alternative, rather intuitive and, in a sense, more elegant approach. There are mainly two reasons why I stick to the Kreps-Wilson treatment: the first reason is that the analysis in Fudenberg and Levine's approach, with its focus on deriving a lower bound for the long-lived player's utility in any Nash equilibrium, tends to leave the exact structure of the underlying behavior unclear. The other reason has to do with the nature of the task at hand: it is not at all clear to me whether their methods can be applied in this environment, as here a long-lived agent by playing his Stackelberg action over and over again cannot hope to eventually convince entrants that this action will be played for sure in the future. At best, he can lead them to expect that this action will be played with probability one half, which equals the probability that an entrant is matched with any given incumbent. This simply because he cannot control the actions of the other long-lived player, and entrants cannot distinguish between incumbents<sup>3</sup>.

Though, as mentioned, this paper works with the notion of reputation originating in the work of Kreps *et al.* (1982), for contrast I should mention the literature that conceptualizes reputation as a "norm" in the context of an infinitely repeated game. Papers following this "norm" approach are Dybvig and Spatt (1980), Klein and Leffler (1981), and Telser (1980). This latter view of reputation is fully forward-looking, and would not seem to capture the backward-looking element in the intuitive, everyday idea of reputation.

A recent paper that claims to study collective reputations is Tirole (1996) (in fact, the only recent contribution on the subject that I am aware of). That paper works with a rather peculiar definition of reputation that does not seem to correspond to either of the two mentioned previously: in Tirole's paper there is neither learning nor repeated strategic interaction (as it studies a random matching game with a continuum of agents). The Tirole contribution strikes me as a dynamic version of Akerlof (1970), in which the dynamics are driven not by strategic considerations, but by the evolution of a physical state (namely, the probability that an agent might be identified as a cheater). In fact, the most recent precursor of this work is Diamond (1989). Even though that paper

---

<sup>2</sup> The results obtained in this paper carry over to a seller/buyer interaction with "fixed prices" in which a seller supplies an item whose quality cannot be ascertained ex-ante by the buyer.

<sup>3</sup> As for introducing richer perturbation sets: why should one work with the more complex model, when, as Fudenberg and Levine have shown, the additional complexity does not really affect the way the model works?

does not deal with the issue of collective reputations per se, it works with an anonymous population which effectively shares a collective reputation. Free-riding is not an issue though, as the population is infinite, and so, individuals decide whether to invest or not in a non-strategic fashion.

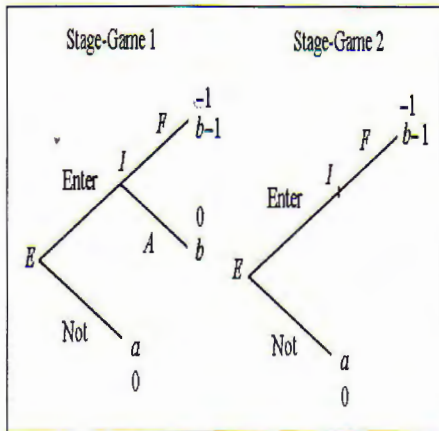
Finally, this work draws on the literature on matching-games and bargaining surveyed in Osborne and Rubinstein (1990), and on the somewhat narrower literature on imperfect information flows in matching games originating in the contributions of Rosenthal (1979), Kandori (1992), and Okuno-Fujiwara and Postlewaite (1995)<sup>4</sup>.

The structure of this paper is as follows. First, the basic matching-game is outlined (Section 3). Then a benchmark case is studied in which incumbents are identifiable. This section basically revisits Kreps and Wilson (1982a) analysis of reputation in the present random matching environment. The following section presents an example of an equilibrium in which there is no reputation building at all due to “free-riding”. The final section shows that, in the  $T = 2$  case, assuming that incumbents know each other types will neutralize such “free-riding”, and allow reputations to develop.

### 3. The Model

The game lasts for  $T$  periods. Following Kreps-Wilson, I will measure time backwards (*i.e.*,  $T$  stands for the first period, and 1 for the last). There are  $2(1 + T)$  players. At any given date, there are 4 players alive:  $2T$ -periods-lived incumbents, denoted by capital  $I$ 's and identified by subscripts 1 or 2; and two one-period-lived entrants, denoted by capital  $E$ 's, identified by their “cohort” affiliation (more on cohorts in a moment). The entrants are matched randomly with the incumbents each period. The incumbent matched with entrant  $E$  will be denoted by  $I(E)$ . Matched agents proceed to play stage-game 1 or 2 as in figure 1.

Figure 1. Stage-Game.



<sup>4</sup> More recent contributions in this last line are Ellison (1994), and Ghosh and Ray (1996).

Incumbents might be of one of two types  $\tau$ : “Rational” ( $R$ ) or “Tough” ( $A$ ). A “rational” incumbent plays stage-game 1, while a “hard-wired” incumbent plays stage-game 2. Note that  $0 < b < 1$ ,  $a > 1$ . Each entrant initially believes that an incumbent is “rational” with probability  $p_T(R)$ . I assume incumbents’ types are drawn independently from identical distributions. I will denote the probability that both incumbents be rational at time  $t$  by  $p_t(R, R)$ . Incumbents are assumed to know whether they are rational or tough. The entrant’s payoffs are as in the diagrams above. The incumbent’s overall payoffs are given by the undiscounted or discounted sum of the payoffs in the stage-games he or she takes part in.

Cohorts of short-lived entrants, of which there will be at most two, in which case they will be identified by superscripts  $F1$  and  $F2$ , are sequences of successive generations of entrants that share their “experiences”. An agent’s “experience” as of time  $t$ , denoted by  $X(A|t)$ , is defined to be all that that agent has observed directly up to and including that period. The objects that might enter an agent’s experience at any given time  $t$ , denoted by  $X_t(A)$ , are given by the actions undertaken at that time in either of the 2 matches, the identities<sup>5</sup> of the agents undertaking those actions, the types (denoted by  $\tau$ ) of the agents alive at  $t$  (if applicable), as well as the overall matching-pattern (who is matched with whom,  $E_t^{F1}/I_1$  or  $E_t^{F1}/I_2$ , denoted by  $MP_t$ ).

In order to distinguish between actions of agents whose identities are assumed to be observable and those of agents who cannot be directly recognized, I will adopt the convention of writing, for example,  $a(E_t^{Fj})$ , for the identified action undertaken by an entrant of cohort  $Fj$  at time  $t$ , while the action of, say, an non-identified entrant belonging to that cohort would be written  $a_t^{E(Fj)}$ . Note that entrants experience will only cover objects existing at the time the entrant is alive, while an incumbent’s experience will contain objects at each date  $t \leq T$  (so will the experience of a cohort). Also, note that the experience of an agent, even a long-lived one, is not the same thing as that agent’s knowledge<sup>6</sup>.

As an example of how this notation works, take the first situation this paper will look at: one where there are two cohorts of entrants, and each period the current generation of entrants includes at most one member of each cohort. Entrants will be assumed not to be able to recognize the incumbents they have interacted with in the past, and only be able to see directly what happens in

---

<sup>5</sup> Note that all agents have identities (by definition), but that does not mean they are recognizable, much less that they are “named”. A “Name” in practice is simply a device that allows a player to associate a specific history with the identity of the player he or she is interacting with. Note that for this identification (*i.e.*, to distinguish a player’s record from that of other players), it is not only necessary that the player being identified carries a name, but also that it be common knowledge among the players that no two players share the same name and that players cannot change their names.

<sup>6</sup> For the limited purposes of this paper, I find this notation more transparent than the more general partitioned formalism conventionally used to describe knowledge.

the matches they take part in. The experience at time  $t$  of a member of cohort  $F^j$ ,  $X_t(E_t^{F^j})$ , is then

$$X_t(E_t^{F^j}) = \left\{ a_t^{I(E_t^{F^j})}, a(E_t^{F^j}) \right\}$$

The experience of the cohort (on the basis of which individual inferences are actually made) is given by:

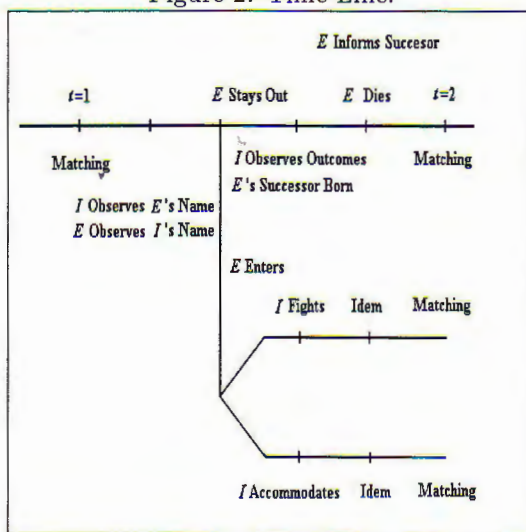
$$FX_t(F^j) = \cup_{t' \geq t} X_{t'}(E_{t'}^{F^j}).$$

The experience as of time  $t$  of an incumbent who is assumed not to be able to observe directly the other incumbent's type (no type-sharing) but who (as always in this paper) is assumed to be able to observe what happens in all matches in the economy, as well as to know the exact matching-pattern at each time and the cohort affiliation of the entrants in each match, satisfies

$$\begin{aligned} X(I_i|t) &= \cup_{t' \geq t} X_{t'}(I_i) \\ &= \cup_{t' \geq t} \left\{ \tau(I_i), a(E_{t'}^{F^j}), a(E_{t'}^{F^i}), a_{t'}(I_1), a_{t'}(I_2), MP_{t'} \right\}. \end{aligned}$$

Let me summarize some of the features all the models dealt with here share. Incumbents know which payoff structure obtains (*i.e.*, they know their type), while entrants do not. Also, I will invariably assume that incumbents can observe each others' actions ex-post; in other words, that at every moment they know the full history of the economy (in other words, there will always be experience-sharing between incumbents). Moreover, I will always assume that all entrants share the same initial beliefs, and that these do not vary across incumbents. Finally, it will be assumed throughout that incumbents can identify the cohort affiliation of entrants. The solution concept I employ is Kreps and Wilson's notion of sequential equilibrium. The following diagram illustrates the time-line of the game:

Figure 2. Time-Line.



#### 4. Anonymity without Free-Riding

Here, I look at the scenario that has all agents identifiable. This is the case “nearest” to Kreps-Wilson, and, as such, a useful (if rather obvious) benchmark.

##### 4.1 Pure Random Matching: A Benchmark

All agents are identifiable and there is no experience-sharing among contemporaneous entrants, nor do incumbents know each other types. In terms of the notation introduced above, the experience of entrants at time  $t$  here can be described as follows

$$X_t \left( E_t^{Fj} \right) = \left\{ a_t(I(E_t^{Fj})), a(E_t^{Fj}), MP_t \right\}$$

The experience of incumbents as of time  $t$  is given by

$$\begin{aligned} X(I_i|t) &= \cup_{t' \geq t} X_{t'}(I_i) \\ &= \cup_{t' \geq t} \left\{ \tau(I_i), a(E_{t'}^{Fj}), a(E_{t'}^{Fi}), a_{t'}(I_1), a_{t'}(I_2), MP_{t'} \right\} \end{aligned}$$

Not surprisingly, the equilibrium of this game strongly resembles that of Kreps and Wilson:

**Proposition 1.** *The equilibrium strategies and beliefs are exactly as in Kreps and Wilson, except that the condition guaranteeing reputation formation (from the period-before-last onward) is now  $(1/2)a > 1$  (rather than  $a > 1$ ), and entrants will mix with probability  $1/2a$  (rather than  $1/a$ ). (For a detailed description, see Appendix).*

*Proof.* The reasoning is, mostly, analogous to Kreps-Wilson’s. Two remarks should suffice to show why this is so: 1) The condition  $(1/2)a > 1$  just states that the maximum expected gain at  $t = 2$  (letting  $t = 1$  be the last period) from deterring entry by fighting,  $(1/2)a$ , exceeds the cost of doing so, 1. If this is so at  $t = 2$ , this must be so as well for any  $t > 2$  (no discounting); 2) Since cohorts are “identifiable” here, the value function of an incumbent can be written as the sum of two independent parts, each giving the expected payoff of interacting with a specific cohort. Moreover, since the incumbents are identified, their value functions will not depend on each others’ actions. This implies that one can think of this game as 4 independent games, one for each cohort/incumbent pairing, and each with payoffs scaled by  $1/2$ . These “subgames” can be analyzed as simple 1:1 games<sup>7</sup>. Applying to each such “subgame” exactly the same arguments as in Kreps-Wilson, and then putting together the resulting “subequilibria”, yields the overall equilibrium described above. ■

<sup>7</sup> An intuitive way of seeing this equivalence is to think of each of these games as one where every period an entrant is matched with an incumbent with probability  $1/2$  or not at all.



Perhaps the only interesting difference between this environment and Kreps-Wilson is contained in the following result:

**Proposition 2.** *The equilibrium described in the previous proposition is unique.*

*Proof.* The on-the-equilibrium-path strategies are unique for the same reasons as in Kreps-Wilson (see Appendix). Overall uniqueness follows from the consistency requirement in the definition of sequential equilibrium<sup>8</sup>, as it can be shown that this requirement fully determines beliefs off-equilibrium-path. The fact that consistency suffices to pin down beliefs off-equilibrium-path is essentially a consequence of working with automata incumbents, instead of incumbents who “like” to fight. ■

Of course, the Kreps-Wilson punchline carries over: As the horizon expands, the critical beliefs’ value separating the entry from the no-entry regions will tend to zero (reputation building).

**Corollary 3.** *As  $T \rightarrow \infty$ , even a very small initial assessment that an incumbent is “hard-wired” will lead to reputation building.*

In conclusion: modifying Kreps-Wilson in this way leaves their results practically unchanged, except for the strengthened uniqueness, and the two minor differences pointed out in Proposition 1. As these latter features are straightforward consequences of assuming that two cohorts of entrants are each being matched randomly with a different incumbent every period, it seems appropriate to refer to them as “pure random matching effects”.

## 5. Free-Riding Under Experience Sharing

This section looks at what happens when there is information-sharing among contemporaneous entrants and incumbents cannot observe each others’ types (the next section looks at the corresponding scenario when incumbents know each others’ types in the case  $T = 2$ ). That is, what happens when all entrants alive at any given time belong to one and the same cohort, while each incumbent ignores the type of the other one.

The section describes a particular class of symmetric equilibria under discounting in which there is no reputation building at all (for a certain parameter constellation). The aim is to illustrate a possibility, rather than provide a complete analysis.

---

<sup>8</sup> That requirement states that an equilibrium system of beliefs must correspond to the limit of a sequence of beliefs’ systems generated via Bayes’ Rule from a sequence of completely mixed strategies which themselves converge to the equilibrium strategies.

### 5.1 Experience Sharing; No Type Sharing

In terms of the notation introduced earlier, this case can be described as follows: The experience at time  $t$  of an entrant  $E_t^j$  is given by

$$X_t(E_t^j) = \left\{ a_t^{I(E_t^j)}, a(E_t^j), a_t^{I(E_t^i)}, a(E_t^i) \right\}$$

The experience of an incumbent  $I_k$  at time  $t$  is

$$X_t(I_k) = \left\{ \tau(I_k), a(E_t^j), a(E_t^i), a(I_k), a(I_l), MP_t \right\}$$

#### 5.1.1 No Reputation Building Regardless of Horizon in the Presence of Discounting

The purpose of this subsection, it bears repeating, is not to present a full analysis of the formation of reputations when there is experience sharing and payoffs are discounted. Rather, the aim is to show how free-riding and discounting can give rise to what seems to me to be a very extreme equilibrium, one in which there are no incentives to invest in reputation regardless of the length of the game (more precisely, in this equilibrium the critical beliefs separating the entry from the no-entry regions remain constant as  $T \rightarrow \infty$ ).

For this kind of result, discounting is necessary. To see this, note that, by the very nature of the result this section is aiming at, it must always be feasible to convince entrants that incumbents are tough in the range of (marginal) beliefs  $[0, b]$ . For, if the critical beliefs are not to shift leftward as the horizon expands, accommodation with probability one must be the prescribed action in equilibrium in this region. But then, by deviating and fighting in this range, there is always a positive probability that an incumbent can convince an entrant that he is tough. In the absence of discounting, the reward from convincing an entrant that one is tough would go to infinity, and it would eventually be impossible to sustain fighting for sure as an equilibrium action.

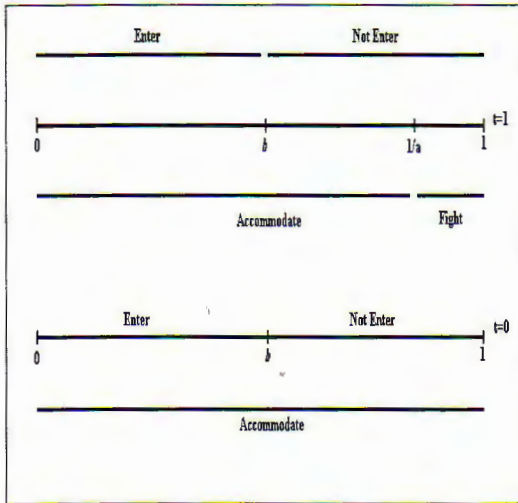
It turns out that it is possible in this equilibrium to use the marginal probability that an incumbent be a compulsive fighter as the state. This because there is no mixing in equilibrium.

**Proposition 4.** *If incumbents cannot observe each others' types, and if  $b > 1/2$ ,  $\beta a > 1$ , and  $\beta a / (1 - \beta) < 1/b$ , then there is no reputation formation even as the horizon goes to infinity.*

*Proof.* See Appendix.

To understand what is going on here, it seems best to start by looking at a two-period, no-discounting version of this game. The diagram below illustrates the discussion:

Figure 3. No Reputation in Two Periods.



In the last period, here, as in all the scenarios in this paper, nothing changes. In the period before last, starting off from initial beliefs somewhere to the left of  $1/a$ , the equilibrium prescribes accommodation with probability 1. To see that accommodation is a best response here, note that if an incumbent deviates from this action and fights instead, her expected payoff will be given by

$$p_2 a + (1 - p_2) 0.$$

For all initial beliefs below  $1/a$ , this expected payoff will not match the additional cost of fighting, namely 1, making the deviation under consideration unprofitable. On the other hand, for initial beliefs in  $[0, 1/a]$ , the net payoff from accommodating is zero. Hence, it is clearly best to do so.

The role of discounting is the following: if incumbents were not to discount, the reward from convincing entrants that incumbents are tough would grow to infinity as the horizon expands, eventually pushing the expected payoff from deviating and fighting above 1. To prevent this, it is necessary that incumbents discount their payoffs at a sufficiently high rate, namely, a rate  $\beta$  such that

$$\beta a / (1 - \beta) < 1/b.$$

The right hand-side of this expression gives the reward from discouraging entry forever, from next period on (hence the additional  $\beta$ ). The condition guarantees that the beliefs making an incumbent indifferent between deviating-and-fighting or conforming-and-accommodating (the inverse of the right hand-side expression), always (*i.e.*, for any horizon) lie to the right of  $b$ . In this way, since the expected gain from deviating varies monotonically with initial beliefs, no

matter how long the game goes on, it will always remain a best response to accommodate in the region  $[0, b]$ .

Note that there is free-riding here because what the other incumbent is doing contemporaneously is modifying the way this incumbent's current actions influence the beliefs of entrants. In fact, it is this feature that makes accommodation with probability one a feasible equilibrium action.

Finally, note that in the 1:1 setup if  $\beta a / (1 - \beta) < 1$ , there will not be reputation building from the start on as well. The reason is simply that fighting is too costly to ever justify investing in reputations. That is not what is going on here: it is being explicitly assumed that  $\beta a > 1$  (which implies that  $\beta a / (1 - \beta) > 1$ )<sup>9</sup>

## 5.2 Reputation Despite Free-Riding

The experience set of an incumbent  $I_k$  at time  $t$  is now given by

$$X_t(I_k) = \left\{ \tau(I_k), \tau(I_l), a(E_t^j), a(E_t^i), a(I_k), a(I_l), MP_t \right\},$$

while the experience of entrants is as in the preceding section.

When incumbents are aware of each others' types, the computation of an equilibrium becomes more involved: Incumbents' strategies depend now on profiles of types, in addition to beliefs and time. Moreover, the formulas for updating beliefs become considerably more complicated, as entrants now have to "average" out all states (profiles of types) under which an observed profile of actions might have arisen.

For these reasons, I study here only the  $T = 2$  case. This allows me to use the marginal probability that an incumbent be tough in describing an equilibrium. Even though the equilibria presented will be recursive in such magnitude<sup>10</sup>, they remain considerably less transparent than those encountered before.

Again, I repeat that the aim of this subsection is to suggest an interesting possibility (that sufficiently widespread information-sharing might neutralize the free-riding problem) rather than provide definitive answers.

**Proposition 5.** *In the case  $T = 2$ , for any given value of the parameter  $b \geq 1/2$ , if incumbents can observe each others' types and actions, entrants share their experiences, and  $a > 1$ , there will be an equilibrium in which reputation-building is only be marginally slowed relative to what it would have been in the 1:1 case. In fact, for a parameter  $b^*$ , there will be an equilibrium with reputation*

<sup>9</sup> All the restrictions on parameters boil down to two conditions:  $b < 1/2$  and  $\beta < 1/2$ .

<sup>10</sup> More precisely, the entrants' strategies will be recursive in this magnitude, while the incumbents' strategies will depend also on the realized profile of types.

*formation exactly as in the Kreps and Wilson model (i.e., reputation formation will not be slowed down at all).*

For a precise description of the equilibrium and a proof of the proposition, the reader is again referred to the Appendix. The first thing to note is that there is still free-riding here. Take the state  $(R, R)$  (i.e., both incumbents are rational), and initial beliefs below  $b$ . Under these conditions, it is equilibrium behavior for an incumbent to accommodate for sure. The reason being that unilateral fighting cannot succeed in convincing an entrant that incumbents are tough. More precisely: The incumbent realizes that should she deviate and fight, the entrant will observe the outcome  $(F, AC)$  (i.e., one incumbent fights; the other accommodates). According to the equilibrium, this outcome should only arise if the state is  $(R, A)$ . Hence, after observing  $(F, AC)$ , the entrant will assign probability  $1/2$  to the event that an incumbent is tough. But this belief is not high enough to deter entry the following period (as  $b > 1/2$ ).

Note further that in the equilibria described, when the state is  $(R, R)$ , incumbents accommodate entry throughout the range of beliefs  $[0, b]$ , while in the Kreps-Wilson game they would have fought entry with positive probability. On the other hand, when the state is  $(R, A)$ , incumbents will, at any given beliefs in this range, fight with a higher probability than they would in the Kreps-Wilson environment. These two circumstances tend to shift the critical beliefs in opposite directions, so that, without further arguments, is hard to say whether the critical beliefs will be above or below those in Kreps-Wilson. But the following diagram makes plain that beliefs will never be above the Kreps-Wilson value.

Interestingly, there will be a value of  $b$ , namely,  $b^*$ , at which there will be no "loss" in reputation relative to Kreps-Wilson. Also, as this parameter takes values near one, the loss in reputation becomes negligible. Although, for any value of  $b$ , the loss in reputation will be relatively small. It would seem that the additional information-sharing among incumbents allows them to better coordinate their responses, and thus avoid the extreme form of free-riding which led to the drastic result of the previous section (fighting can now be sustained when the state is  $(R, A)$ , precisely because the rational incumbent can now be sure that the other incumbent will not accommodate).

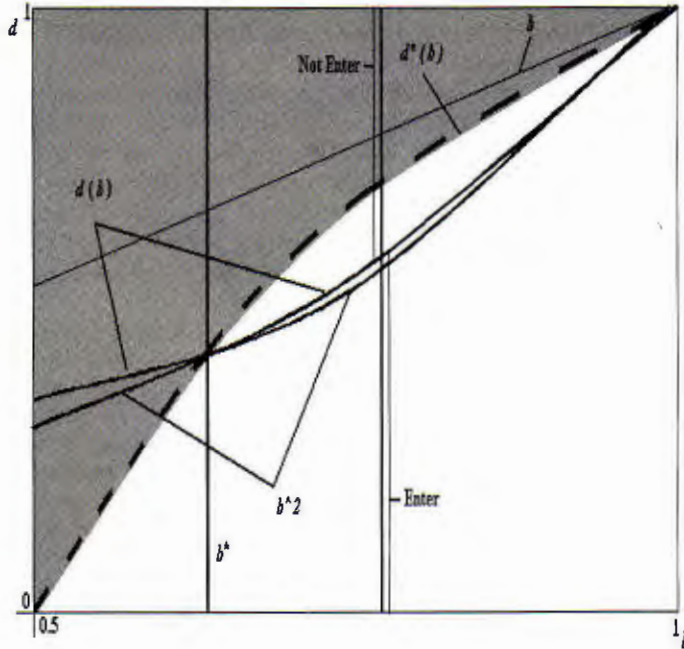
The issue of uniqueness is left open. Note that it is not even clear whether the equilibrium presented here is unique among the class of recursive equilibria.

## 6. Final Remarks

The two main results of this paper (Propositions 4 and 5) show that it might be misleading to analyze reputation formation in one-on-one environments since the exact pattern of information-sharing will be crucial in molding incentives to invest in a reputation in collective setups as the one considered here. Experience-sharing among contemporaneous entrants in the absence of type-sharing generates free-riding, but type-sharing accompanied by experience-sharing would seem to neutralize the adverse effects of the resulting free-riding.

Of course, what I have done here is to illustrate possibilities. A more general analysis is outstanding.

Figure 4. Equilibria under Full Information-Sharing.



## Appendix

### A.1 Description of Equilibrium for Proposition 1

For  $(1/2)a > 1$ , the following strategies and beliefs form a sequential equilibrium of the matching game just described:

Starting from given initial beliefs  $p_T$ , define beliefs recursively as follows:

- a) If there is no entry at stage  $t + 1$ , then  $p_t = p_{t+1}$ .
- b) If there is entry at stage  $t + 1$ , this entry is fought, and  $p_{t+1} > 0$ , then  $p_t = \max(b^n, p_{t+1})$ .
- c) If there is entry at stage  $t+1$ , and either this entry is met by accommodation or  $p_{t+1} = 0$ , then  $p_t = 0$ .

The rational incumbent's strategy is given by:

If  $t = 1$ , the incumbent accommodates. If  $t > 1$  and  $p_t \geq b^{t-1}$ , the incumbent fights. If  $t > 1$  and  $p_t < b^{t-1}$ , the incumbent fights with probability  $(1 - b^{t-1}) p_t / (1 - p_t) b^{t-1}$ , and accommodates with the complementary probability.

The entrant's strategy is given by:

If  $p_t > b$ , entrant stays out. If  $p_t < b$ , then entrant enters. If  $p_t = b^t$ , then entrant stays out with probability  $(1/2)a$ .

### A.2 Proof of Proposition 4

I present the argument for the two period case. Since the equilibrium is stationary, generalizing the result to longer horizons is straightforward. This is the equilibrium (with the last period corresponding to  $t = 0$ ):

Beliefs:

- a) If no entry takes place or  $p_{t+1} = 0$ , then  $p_t = p_{t+1}$ .
- b) If  $t \geq 1$ ,

$$0 < p_{t+1} \leq \left[ \frac{1 - \beta^t}{1 - \beta} \beta a \right]^{-1},$$

and the observed action profile is:

- i)  $(F, F)$ , then  $p_t = 1$ .
- ii)  $(F, AC)$ , then  $p_t = 1/2$ .
- iii)  $(AC, AC)$ , then  $p_t = 0$ .

Strategies:

Incumbents:

- a) If  $t \geq 1$  and

$$p_{t+1} < \left[ \frac{1 - \beta^t}{1 - \beta} \beta a \right]^{-1},$$

then accommodate.

- b) If  $t \geq 1$  and

$$p_{t+1} \geq \left[ \frac{1 - \beta^t}{1 - \beta} \beta a \right]^{-1},$$

then fight.

If  $t = 0$ , accommodate always.

Entrants:

If  $p_{t+1} < b$ , then enter; if  $p_{t+1} \geq b$ , stay out.

### A.2.1 Optimality of Strategies

Incumbent:

- I)  $\delta_2 < 1/(\beta a)$ , if the incumbent fights instead of accommodating, then, with probability  $\delta_2$ , the other incumbent will be tough. In that case, the entrant will observe outcome  $(F, F)$ , and, hence,  $\delta_1 = 1$ , thus deterring entry. With probability  $1 - \delta_2$ , on the other hand, the outcome will be  $(F, AC)$ , and entry will not be deterred. So, the expected payoff of fighting in this beliefs' region is given by

$$\delta_2 \beta a - 1 \leq 0$$

In other words, it is best to accommodate.

- II)  $\delta_2 \geq (\beta a)^{-1}$ , now it is best to fight as accommodation will surely induce entry, while fighting lead to an outcome of  $(F, F)$ , and, so, to beliefs  $\delta_1 = \delta_2 \geq (\beta a)^{-1} \geq b$ , with equality iff  $\delta_2 = (\beta a)^{-1}$ .

Entrant:

Since in the region  $[0, (\beta a)^{-1}]$  at  $T = 2$ , an incumbent fights only if she is tough, the critical value separating the entry from the no-entry regions,  $\delta_2(b)$ , is given by the solution to the following equation,

$$p(F|\delta_2 \in [0, 1/a]) (b - 1) + (1 - p(F|\delta_2 \in [0, 1/a])) b = 0$$

Since  $p(F|\delta_2 \in [0, 1/a]) = \delta_2^2 + 2\delta_2(1 - \delta_2)(1/2)$ , this yields  $\delta_2(b) = b$ .

### A.2.2 Consistency of Beliefs

- I)  $\delta_2 < 1/(\beta a)$  :

- i)* If observed outcome is  $(F, F)$  : Since rational incumbents are supposed to accommodate in that region, the entrant should conclude  $\delta_1 = 1$ , *i.e.*, that the incumbent is tough for sure.
- ii)* If  $(AC, AC)$ , then, evidently,  $\delta_1 = 0$ .

- iii)* If  $(F, AC)$ , then  $\delta_1 = 1/2$ , as only a tough incumbent can be expected to fight in this region.

- II)  $\delta_2 \geq 1/(\beta a)$  :

- i)* If  $(F, F)$ , then as both rational and tough types are supposed to fight, priors should equal posteriors.
- ii)* If  $(AC, AC)$ , then evidently  $\delta_1 = 0$ .
- iii)* If  $(F, AC)$  : It follows that  $\delta_1 = (1/2)\delta_2$  as now both tough and rational incumbents are supposed to fight, and deviations are taken to be uncoordinated. This is an out-of-equilibrium outcome, yet the consistency requirement in the definition of sequential equilibrium suffices to pin down beliefs. Notice also that this equilibrium is not unique.



There is another equilibrium that has incumbents fighting with probability 1 for beliefs above  $b$ , for all  $t$ . The path of play, though, is the same in both equilibria. Besides there are yet other equilibria in which there is mixing in the region  $[0, b]$ .

### A.3 Proof of Proposition 5

The equilibrium assessment is as follows:

Beliefs:

- a) If no entry takes place or  $p_{t+1} = 0$ , then  $p_t = p_{t+1}$ .
- b) If  $0 < p_{t+1} \leq \bar{\delta}(b)$  and observed action profile is:
  - i)  $(F, F)$ , then  $p_t = b$ .
  - ii)  $(F, AC)$ , then  $p_t = 1/2$ .
  - iii)  $(AC, AC)$ , then  $p_t = 0$ .
- c) If  $\bar{\delta}(b) < p_{t+1} \leq b$  and observed action profile is:
  - i)  $(F, F)$ , then  $p_t \geq b$ .
  - ii)  $(F, AC)$ , then  $p_t \leq 1/2$ .
  - iii)  $(AC, AC)$ , then  $p_t = 0$ .
- d) If  $p_{t+1} \geq b$  and observed action profile is:
  - i)  $(F, F)$ , then  $p_t = p_{t+1}$ .
  - ii)  $(F, AC)$ , then  $p_t = (1/2)p_{t+1}$ .
  - iii)  $(AC, AC)$ , then  $p_t = 0$ .

Strategies:

Incumbents:

If realized types are  $(R, R)$  and  $t > 0$ :

- a) If  $p_{t+1} < b$ , then accommodate.
- b) If  $p_{t+1} \geq b$ , then fight.

If realized types are  $(R, A)$  and  $t > 0$ :

- a) If  $p_{t+1} < \bar{\delta}(b)$ , then mix, fighting with probability

$$\frac{(1-b)\bar{\delta}}{(2b-1)(1-\bar{\delta})},$$

where  $\bar{\delta}$  stands for initial beliefs.

- b) If  $p_{t+1} \geq \bar{\delta}(b)$ , then fight.

If  $t = 0$ , accommodate always.

Entrants:

At  $t \neq 0$ , if  $p_{t+1} < \delta(b)$ , then enter; if  $p_{t+1} \geq \delta(b)$ , stay out. At  $t = 1$ , if  $p_{t+1} < b$ , enter; if  $p_{t+1} > b$ , stay out; if  $p_{t+1} = b$  then mix, staying out with probability  $1/a$ . In this case,

$$\delta(b) = \begin{cases} \frac{1 - 2b + \sqrt{-8b^3 + 16b^2 - 8b + 1}}{2(1 - b)}, & \text{if } b \geq b^*, \\ 1 - (1 - b)^{1/2}, & \text{if } b < b^*, \end{cases}$$

where  $b^*$  is subject to  $1 - (1 - b)^{1/2} = \tilde{\delta}(b)$  and  $\tilde{\delta}(b) = (2b - 1)/b$ . To prove that the candidate equilibrium is in fact one, I check the optimality of strategies and the consistency of beliefs:

### A.3.1 Optimality of Strategies

Incumbents:

I) If state is  $(R, R)$ :

- i)  $\delta_2 < b$ , best to accommodate, since if, instead, incumbent fights, resulting beliefs will be given by  $\delta_1 = 1/2 < b$ , and so entry will not be deterred.
- ii)  $\delta_2 \geq b$ , best to fight, since if, instead, the incumbent accommodates, the resulting beliefs will be given by  $\delta_1 = (1/2)\delta_2 < 1/2 < b$  (observed outcome would be  $(F, AC)$ ). If fight, then beliefs would be given by  $\delta_1 = \delta_2 \geq b$ , so incumbent must be either indifferent between fighting and accommodating, or strictly prefer the former to the latter.

II) If state is  $(R, A)$ :

- i)  $\delta_2 < \tilde{\delta}(b)$ , fighting with probability 1 will lead to  $\delta_1 < b$  after  $(F, F)$  is observed. To see this: Note that the probability of an entrant being currently matched with an incumbent who is of type  $A$ , after observing  $(F, F)$ , is given by

$$\begin{aligned} p(A | F, F) &= p((A, A) | (F, F)) + \frac{1}{2}p((A, R) | (F, F)) \\ &\quad + \frac{1}{2}p((R, A) | (F, F)) + 0p((R, R) | (F, F)) \\ &= \frac{\delta_2^2}{p(F, F)} + \frac{p(F | R, (R, A))\delta_2(1 - \delta_2)}{p(F, F)}, \end{aligned}$$

with  $p(F | R, (R, A)) = p((F, F) | (A, R)) \equiv x$ . Given that there is accommodation by  $R$  when the state is  $(R, R)$ , we have

$$\begin{aligned} p(F, F) &= \delta_2^2 + 2x\delta_2(1 - \delta_2), \\ p(A | (F, F)) &= \frac{\delta_2^2 + x\delta_2(1 - \delta_2)}{\delta_2^2 + 2x\delta_2(1 - \delta_2)} \end{aligned}$$

Setting  $p(A | (F, F)) = b$ , solving for  $x$  satisfying this equation, and denoting such an  $x$  by  $\tilde{x}$ , we get

$$\tilde{x}(b) = \frac{(1-b)\delta_2}{(1-\delta_2)(2b-1)}.$$

Note that  $\tilde{x}(b)$  is increasing in  $\delta_2$ . Now, define  $\tilde{\delta}$  by setting

$$\frac{(1-b)\tilde{\delta}}{(1-\tilde{\delta})(2b-1)} = 1,$$

to get

$$\tilde{\delta}(b) = \frac{2b-1}{b}.$$

All  $\delta_2 > \tilde{\delta}(b)$  will yield  $x > 1$ ; all  $\delta_2 \leq \tilde{\delta}(b)$  will result in  $x < 1$ . For  $\delta_2 < \tilde{\delta}(b)$ ,  $\tilde{x} < 1$ , and so, if a rational incumbent fights with probability 1, we get  $x > \tilde{x}$ , and, hence,  $p(A | (F, F)) < b$  (since  $p(A | (F, F))$  is decreasing in  $x$ ). It follows that it does not pay to fight. If, instead, the incumbent accommodates with probability 1, this would imply that, when  $\delta_2 < \tilde{\delta}(b)$ , incumbents accommodate across all states ( $\tilde{\delta}(b) \leq b$  for  $b \in [\frac{1}{2}, 1]$ ), and there would be an incentive to fight instead when the state is  $(R, A)$ , as the observed outcome would be  $(F, F)$ , and, so,  $\delta_1 = 1$ . It follows that incumbents must mix. In order for the incumbent to be indifferent between fighting and accommodating, it must be that an entrant, after observing  $(F, F)$  when  $\delta_2 < \tilde{\delta}(b)$ , must himself mix, so that

$$p(E | (F, F), (\delta_2 < \tilde{\delta}(b)))0 + (1 - p(E | (F, F), (\delta_2 < \tilde{\delta}(b))))a - 1 = 0,$$

$$\Rightarrow p(E | (F, F), (\delta_2 < \tilde{\delta}(b))) = 1 - \frac{1}{a}.$$

In order to make the entrant willing to mix after observing  $(F, F)$  in the region  $\delta_2 < \tilde{\delta}(b)$ , it must be that  $p(A | (F, F)) = b$ . This implies that incumbents must fight with probability

$$\tilde{x}(b) = \frac{(1-b)\delta_2}{(1-\delta_2)(2b-1)}.$$

- ii)  $\tilde{\delta}(b) \leq \delta_2 < b$ , if you accommodate, the observed outcome will be  $(F, AC)$ , and, hence,  $\delta_1 \leq 1/2$ . Since  $b > 1/2$ , entry takes place. If, instead, the incumbent fights, the observed would be  $(F, F)$ , and the resulting beliefs  $\delta_1 \geq b$ . Hence, entry would be deterred. To see this: To obtain  $p(A | (F, F)) = b$ , incumbent has to fight with probability  $\tilde{x}(b)$ . But whenever  $\tilde{\delta}(b) \leq \delta_2$ ,  $\tilde{x}(b) > 1$ . Since  $p(A | (F, F))$  is falling in  $x$ , it must be that

$$p(A | (F, F))(1) \geq p(A | (F, F))(\tilde{x}(b)) = b$$

(with equality only if  $\delta_2 = \tilde{\delta}(b)$ ).

- iii)  $\delta_2 \geq b$ , accommodating leads to observation  $(AC, F)$ , and beliefs  $\delta_1 = (1/2)\delta_2 < 1/2 < b$ . Hence, entry is not deterred. Fighting, on the other hand, leads to observation  $(F, F)$ , and beliefs  $\delta_1 = \delta_2 > b$ . Hence deterring entry.

Entrant:

At  $T = 2$ , in the region  $[0, b]$ , the probability that an entrant is fought (across all states) is given by the solution to the following equation (which must hold if the expected value of entry is to be 0):

$$\delta_2^2 + 2(1 - \delta_2)\delta_2\left(\frac{1}{2}x + \frac{1}{2}\right) = b.$$

Substituting

$$\frac{(1 - b)\delta}{(1 - \delta)(2b - 1)}$$

for  $x$  above, solving for  $\delta_2$ , and picking the positive root, one obtains

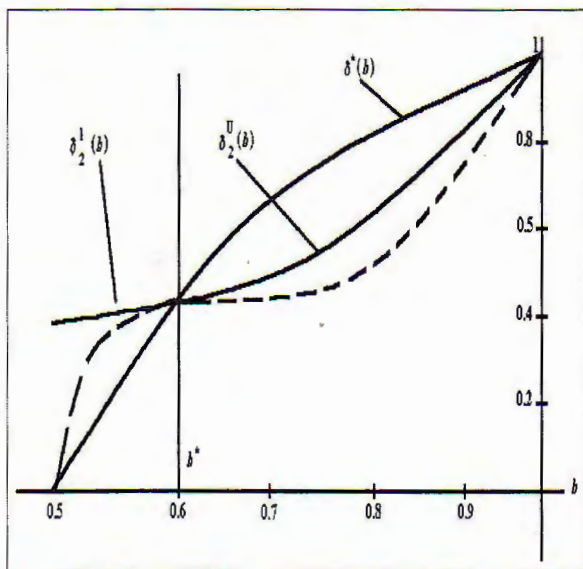
$$\delta_2^U(b) = \frac{(1 - 2b) + (16b^2 - 8b^3 - 8b + 1)^{\frac{1}{2}}}{2(b - 1)}$$

Substituting instead  $x = 1$ , and solving, one obtains

$$\delta_2^1(b) = 1 - (1 - b)^{\frac{1}{2}}$$

The following plot illustrates the properties of these functions:

Figure 5. Properties of Functions.



From the diagram, note that in the region  $[\frac{1}{2}, b^*]$ ,  $\delta_2^U(b) > \tilde{\delta}(b) \Rightarrow \tilde{x}(b) > 1$  (as  $\tilde{x}(b)$  is increasing in  $\delta$ ). It follows that  $\delta_2^U(b)$  overestimates the probability of fighting. Note that the probability that an entrant is fought,

$$\delta^2 + (1 - \delta) \delta \left( \frac{(1 - b) \delta}{(1 - \delta)(2b - 1)} + 1 \right)$$

is increasing in  $\delta$ . Also,  $\delta^2 + 2(1 - \delta) \delta$  is increasing in  $\delta$ . It follows that the true critical value of  $\delta$  is given by  $\delta_2^1(b)$ , for  $[0.5, b^*]$ ; and by  $\delta_2^U(b)$ , for  $[b^*, 1]$ . Evidently, for  $\delta < \delta(b)$ , entry should take place; not so for  $\delta > \delta(b)$  (this follows from the payoffs to entry being monotonically decreasing as beliefs increase since the probability of being fought is increasing, as just shown).

### A.3.2 Consistency of Beliefs

i)  $\delta_2 < \tilde{\delta}(b)$  :

- i) if observation is  $(F, F)$ , by construction.
- ii) If  $(F, AC)$ , if the state were  $(R, R)$ , there would be accommodation (note that  $\tilde{\delta}(b) < b$ ; see plot above). It must be that the state is  $(R, A)$  when this observation is made.
- iii) If  $(AC, AC)$ , obvious.

II)  $\tilde{\delta}(b) \leq \delta_2 < b$ :

- i) If  $(F, AC)$ , at least one incumbent is rational, and so posterior beliefs must be below or at  $1/2$ .
- ii) If  $(F, F)$ , posterior beliefs must be above or at  $b$  by a previous argument.
- iii) If  $(AC, AC)$ , obvious.

III)  $\delta_2 > b$ :

- i) If  $(F, F)$ , since incumbent fights regardless of type,  $\delta_2 = \delta_1$ .
- ii) If  $(F, AC)$ , since someone deviated by accommodating, it must be that  $\delta_1 \leq 1/2$ . From the definition of sequential equilibrium (which implies that deviations should be uncoordinated), it follows that  $\delta_1 = (1/2)\delta_2$ .
- iii) If  $(AC, AC)$ , obvious.

## References

- Akerlof, G. (1970). The Market for Lemons. *Quarterly Journal of Economics*, 84, pp. 488-500.
- Corbin, A. (1987-1991). The Secret of the Individual. In Ari P. and Georges Duby (Eds.). A History of Private Life. Belknap, Harvard University Press, Cambridge, Mass., 5 vols.
- Brickley, J. (1996). Incentives Conflicts and Contracting: Evidence from Franchising. Mimeo, University of Rochester.
- Brickley, J., and R. Dark (1987). The Choice of Organizational Form: The Case of Franchising. *Journal of Financial Economics*, 18, pp. 101-132.
- Diamond, D. W. (1989). Reputation Acquisition in Debt Markets. *Journal of Political Economy*, 97, pp. 828-862.
- Dybvig, P. H., and C. S. Spatt (1980). Does it Pay to Maintain a Reputation?. Mimeo.
- Ellison, G. (1994). Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. *Review of Economic Studies*, 61, pp. 567-588.
- Fudenberg, D., and D. Levine (1989). Reputation and Equilibrium Selection in Games with a Patient Player. *Econometrica*, 57, pp. 759-778.
- Fudenberg, D., and D. Levine (1992). Maintaining a Reputation when Strategies are Imperfectly Observed. *Review of Economic Studies*, 59, pp. 561-579.
- Fudenberg, D., and J. J. Tirole (1992). Game Theory. The MIT Press, Cambridge, Mass.
- Ghosh, P., and D. Ray (1996). Cooperation in Community Interaction Without Information Flows. *Review of Economic Studies*, 63, pp. 491-519.
- Kandori, M. (1992). Social Norms and Community Enforcement. *Review of Economic Studies*, 59, pp. 63-80.
- Klein, B., and K. B. Leffler (1981). The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy*, 89, pp. 615-641.
- Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982). Rational Cooperation in the Finitely Repeated Prisoners' Dilemma. *Journal of Economic Theory*, 27, pp. 245-252.
- Kreps, D. M., and R. Wilson (1982a). Reputation and Imperfect Information. *Journal of Economic Theory*, 27, pp. 253-279.
- Kreps, D. M., and R. Wilson (1982). Sequential Equilibria. *Econometrica*, 50, pp. 862-894.
- Mathewson, G., and R. Winter (1984). The Economics of Franchise Contracts. *Journal of Law and Economics*, 28, pp. 503-526.
- Milgrom, P., and J. Roberts (1982a). Predation, Reputation, and Entry Deterrence. *Journal of Economic Theory*, 27, pp. 280-312.

- Milgrom, P., and J. Roberts (1982). Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis. *Econometrica*, 50, pp. 443-459.
- Okuno-Fujiwara, M., and A. Postlewaite (1995). Social Norms and Random Matching. *Games and Economic Behavior*, 9, pp.79-109.
- Osborne, M. J., and A. Rubinstein (1990). *Bargaining and Markets*. Academic Press, San Diego, Cal.
- Rubin, P. (1978). The Theory of the Firm and the Structure of the Franchise Contract. *Journal of Law and Economics*, 21, pp. 223-233.
- Rosenthal, R. (1979). Sequences of Games with Varying Opponents. *Econometrica*, 49, pp. 1353-1366.
- Strasser, S. (1989). *Satisfaction Guaranteed*. Pantheon Books, New York.
- Telser, L. A. (1980). Theory of Self-Enforcing Agreements. *Journal of Business*, 53, pp. 27-44.
- Tirole, J. (1996). A Theory of Collective Reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, 63, pp. 1-22.
- Wilson, R. (1985). Reputations in Games and Markets. In Roth A. E. (Ed.). *Game Theoretic Models of Bargaining*. Cambridge University Press, New York, pp. 27-62.