Discussion Paper No. 1009

Reputation in Dynamic Games

by

Marco Celentani*
Universidad Carlos III de Madrid
and
Wolfgang Pesendorfer
Northwestern University

October 1992
Revised: April 1993

* Marco Celentani gratefully acknowledges the financial support of the National Science Foundation, Grant SES 90-23697. Wolfgang Pesendorfer gratefully acknowledges the financial support of the National Science Foundation, Grant SES 90-7999. We wish to thank V.V. Chari, David Levine, Andy Newman, Asher Wolinsky and seminar participants at Northwestern University, Universidad Carlos III de Madrid, Universitat Pompeu Fabre and UCLA for useful discussions and suggestions. All the remaining errors are ours.

Abstract

We consider an infinite dynamic game between a large player and a large number (continuum) of small players in which state variables affect players' payoffs. We analyze the perturbed game in which with small probability the large player can be one of many commitment types. We show that this perturbation allows the large player to establish a "reputation" and that in every Nash equilibrium the large player gets at least what he could get by committing to an optimal strategy as his discount factor approaches 1. Furthermore, if the transition function is *reversible*, i.e. if a small player can move from one state to another only if he can also return, then this result holds even if the small players' discount factor also approaches 1.

We present an example of a firm associated with a large number of employees who can accumulate firm specific human capital. A tight bound on the firm's equilibrium payoff is given and it is shown how reputation can allow the firm to improve upon *all* subgame perfect equilibria of the unperturbed game. For the durable goods monopoly it is shown that reputation may fail to improve the monopolist's payoff.

KEYWORDS: Dynamic games, reputation, Nash equilibrium.

Journal of Economic Literature Classification Numbers: C72, C73.

1 Introduction

Many economic problems have the feature that a *state variable* such as capital, debt or money, provides a link between present actions and future payoff opportunities. As an example, games that describe the strategic interaction between a government and households usually involve state variables. It is in this context that the problem of time consistency of optimal government policy arises: since ex-ante and ex-post optimal policies differ, even a benevolent government may not be able to achieve the optimal commitment outcome.

Recent work has turned attention to this kind of games. Dutta (1991) provides a Folk theorem for stochastic games. Chari and Kehoe (1990) and Stokey (1992) study the time inconsistency problem introduced by Kydland and Prescott (1977, 1980) and Fischer (1980) and characterize the set of equilibria in problems of optimal policy design when the government cannot commit. Both Chari and Kehoe (1990) and Stokey (1992) show that, if there is sufficiently little discounting, a desirable outcome (the Ramsey outcome in a capital taxation model, Ramsey (1927)), can arise in equilibrium. However, in their model the Ramsey outcome is only one of many equilibria.[1]

We consider a general class of dynamic games with one large player and a large number of small players. A deterministic transition law describes the evolution of the state variable. The large player has some private information about his type, i.e. the small players are uncertain about the type of large player they are facing. This uncertainty may be very small in the sense that the large player is of one particular type with a probability close to one. The goal of this paper is to find conditions under which a patient large player can exploit the uncertainty of his opponents and enforce an outcome that is essentially equivalent to publicly committing to an optimal strategy.

The introduction of uncertainty relative to the type of a player and the consequent possibility of acquiring a reputation for an appropriate behavior has received considerable attention in the literature. Starting with the work of Kreps and Wilson (1982) and Milgrom and Roberts (1982) the studies of reputation effects have focused exclusively on repeated games.

Fudenberg and Levine (1989) study a class of repeated games in which a long lived player faces a sequence of short lived opponents, each of whom plays only once but observes

[1]It is sometimes argued that in this case the government may be able to select its preferred equilibrium. However, Dekel and Farell (1990) show that these selection arguments are inconsistent.

the entire history of the game. If there is a positive probability that the long lived player is a type who always plays the strategy to which the normal player would like to commit, then reputation effects lead to a sharp prediction for all Nash equilibria of the game: the large player will receive a payoff that is at least as large as what he would receive if he could publicly commit to his preferred strategy.

This result is robust in the sense that it does not rely on a refinement of Nash equilibrium and that it is unaffected by further perturbations of the information structure of the game, i.e. by the introduction of additional commitment types[2].

The present paper uses reputational arguments and provides conditions under which results analogous to the ones obtained by Fudenberg and Levine (1989) apply to dynamic games and also provides conditions under which reputational arguments may fail.

Since we consider games with a a large number (continuum) of small players, we will assume that the individual play of the small players is not observed. In a purely repeated game this assumption would imply that each small player behaves like a short-lived player, since his actions will affect neither his future payoffs nor the public history of the game. In a dynamic game the presence of state variables creates an intertemporal link and introduces a new strategic dimension to the problem. Even though a small player cannot influence his opponent's future play, he can change the value of his individual state variable, thereby affecting his own future payoff opportunities. Therefore a small player's behavior will depend on the (expected) future actions of the large player [3].

For example, in a capital taxation model in order to choose a high investment level today, the households in the economy need to become convinced that the government will set low capital tax rates not only today but also in the future.

As is clear, the presence of state variables makes it more difficult for the large player to establish a reputation: small players have to become convinced that the large player will follow a particular strategy not only in the current period but also in the future. The more the small players' behavior is affected by play in the distant future the harder it will be for the large player to gain from establishing a reputation.

Our first result (Theorem 1) applies to the case where the small players have a fixed discount factor while the large player is arbitrarily patient. If there is a commitment type

[2] For extensions of Fudenberg and Levine (1989) see Fudenberg and Levine (1992), Schmidt (1991) and Cripps and Thomas (1992).

[3] See also Schmidt (1992), for a similar effect in games with two long run players.

that plays the strategy to which the large player would want to commit then in any Nash equilibrium the large player is guaranteed at least the optimal commitment payoff[4].

To obtain a result that holds for a wide range of interesting economic applications we allow the payoffs to the small players to depend on the aggregate play of the small players and the aggregate state variable, as well as on their own play and the play of the large player. In the terminology to be introduced, we allow for *strategic externalities* among small players. This has the surprising implication that, for a fixed discount factor, arbitrarily distant play of the large player may affect current behavior of the small players (see Example 2). If the optimal commitment strategy can be approximated by an *eventually periodic* sequence, i.e. a sequence that converges to some cycle of bounded length in finitely many periods, then also in this case reputation effects allow a precise characterization of equilibria. Assuming that the discount factor of the small players stays fixed while the large player gets arbitrarily patient, the large player will receive at least the optimal commitment payoff in all Nash equilibria (Theorem 2).

Finally, we consider the case where both the large and the small players are arbitrarily patient. This case is particularly relevant for policy games, in which, for example, the payoff function of the government is equal to the payoff function of the median voter. Then, if players are very patient, the small players' action may be affected by very distant future outcomes.

In this case it is shown (Theorem 3) that the large player will only be able to exploit his reputation if the following reversibility condition on the transition function is satisfied. A transition function is *reversible* if players can move from one state to another only if they can also return. This condition is satisfied in capital accumulation games, but is not satisfied, for example, in the standard durable goods monopoly [5]. Once a customer has purchased the durable good, he has reached an irreversible state. Example 3 shows how in the durable goods monopoly reputational arguments fail to guarantee the large player his optimal commitment payoff.

An example that is discussed in the paper is that of a firm with a large number of employees who can invest in firm specific human capital (skills) to increases their productivity.

[4]By the optimal commitment payoff, we mean the maximal time average that the large player could guarantee himself by publicly precommitting if the game started in the worst possible state from the large player's point of view.

[5]See for example Coase (1972), Ausubel and Deneckere (1989), Stokey (1981), Fudenberg, Levine and Tirole (1985), Gul, Sonnenschein and Wilson (1986).

Employees are not willing to invest unless the firm rewards the increased productivity. Since human capital is firm specific, however, the firm has an incentive to renege on its promise to reward increased productivity and to pay only the reservation wage. Even though in the game with complete information there may be no subgame perfect equilibrium in which human capital is accumulated (Proposition 1), the introduction of a small amount of uncertainty about the type of the firm implies that the optimal level of human capital will be accumulated and the firm will appropriate all the net gains from this accumulation (Proposition 2).

An interesting application of the case in which large and small players have the same discount factor is the classical time inconsistency problem in an intertemporal capital taxation model (Kydland and Prescott, 1977). Fischer (1980) describes a situation in which a benevolent government has to finance a public good by levying taxes on capital and labor. If the government could commit to a certain strategy, it could achieve the Ramsey outcome (Ramsey, 1927), i.e. the sequence of combinations of capital and labor tax rates that minimize distortions. Once capital has been accumulated, however, it is optimal for the government to raise as much revenue as possible from capital taxation that is ex-post non distortionary. If private investors expect the government to renege on its promise to set low capital tax rates, they will accumulate a suboptimal level of capital.

The result of the present paper is that if the prior probability of a particular commitment type is strictly positive, then in any Nash equilibrium the government will achieve a payoff close to the payoff corresponding to credible commitment to an optimal tax rate.

The structure of the paper is as follows. In section 2 we describe the complete information game. Section 3 introduces the perturbed game, i.e. the possibility of the large player to be one of many "types". Section 4 provides the first result for the case where the discount factor of the small players stays fixed while the large player is very patient. Section 5 discusses the example of firm specific human capital accumulation referred to above. Section 6 gives the result for the case in which there are strategic externalities. Section 7 deals with the case where large and small players share a common discount factor and Section 8 provides conclusions. Proofs are presented in Section 9.

2 Description of the Game

The class of games we consider has two types of players: one large player denoted by b, and a continuum of identical small players $i \in [0,1] = I$. The finite sets Y and X denote the actions of the large player and the small players respectively; $y \in Y$, $x \in X$. Furthermore we let Σ denote the set of mixed actions of the large player.

Each small player has an individual state variable, $z \in Z$, where Z is the state space that is assumed to be finite and identical for all small players. Let Λ denote the set of probability measure on Z, $\lambda \in \Lambda$, and M denote the set of probability measures on $Z \times X$, $\mu \in M$; $\mu(z,x)$ is to be interpreted as the measure of small players with initial value of the state variable equal to z that choose action x. Finally let $\mu^Z \in \Lambda$ denote the marginal of μ on Z and μ^X the marginal of μ on X (μ^X belongs to the set of probability measures on X).

The game is played in the following way: At the beginning of each period $t = 1, 2, \ldots$ the public history (to be described below) is observed by all players and each small player observes his own private history. Conditional on these observations, every small player takes an action $x^i \in X$ and the large player simultaneously takes a (possibly mixed) action $\sigma \in \Sigma$, where Σ denotes the set of probability distributions on Y.

After these actions have been selected, payoffs occur and all players observe the realization of the action of the large player y_t and the distribution μ_t of actions of the small players. Note that this is a joint distribution over actions and states, i.e. after each period every player knows which proportion of players in state z played action x, for every $z \in Z$. Clearly this joint distribution has to be consistent with the state in the beginning of the period $\lambda_t \in \Lambda$, i.e. $\mu_t^Z = \lambda_t$.

The law of motion for the individual state is described by the following function:

$$f : Y \times X \times Z \to Z$$

i.e. $z_{t+1}^i = f(y_t, x_t^i, z_t^i)$. In other words we assume that the value of the individual state variable at date $t+1$ does not depend on the aggregate distribution of the state variable or on the aggregate action played by the small players.

The aggregate law of motion is described by:

$$F : Y \times M \to \Lambda$$

where

$$F(y_t,\mu_t)(z) = \sum_{\{(x,z')|f(y_t,x,z')=z\}} \mu_t(x,z')$$

Note that F is continuous.

Let the distance between μ_t and μ_t' be defined as

$$|\mu_t - \mu_t'| = \sum_{Z\times X} |\mu_t(x,z) - \mu_t'(x,z)|$$

and let

$$|\lambda - \lambda'| = \sum_{Z} |\lambda(z) - \lambda'(z)|.$$

A public history of the game at time t is the sequence of realizations of $y_{t'}, t' = 1,\ldots,t-1$, $\mu_{t'}, t' = 1,\ldots,t-1$ and the aggregate state in period t, λ_t. Since we will want to use a recursive definition of histories we also include λ_τ $\tau = 1,\ldots,t-1$ in the history at time t[6]. The set of histories in period t is denoted by $H_t = (Y \times M \times \Lambda)^{t-1} \times \Lambda$, with $h_t \in H_t$; $h_1 = \lambda_1$ and $h_t = (h_{t-1}; y_{t-1}, \mu_{t-1}, \lambda_t)$ for $t > 1$; $H = H_\infty$. For the history from t' to t, $t' \leq t$ we write $h_t \backslash h_{t'} \in H_{t-t'}$.

For a given sequence of play $(y,\mu) = ((y_1,y_2,\ldots),(\mu_1,\mu_2,\ldots))$ the payoff to the large player is:

$$V^b(\beta,y,\mu) = (1-\beta)\sum_{t=1}^{\infty} \beta^{t-1} v^b(y_t,\mu_t)$$

Similarly for a given sequence $(y,\mu,x,z) = ((y_t,\mu_t,x_t,z_t)_{t=1}^{\infty})$ the payoff to a small player is:

$$V(\delta,y,\mu,x,z) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v(y_t,\mu_t,x_t,z_t)$$

Since the small players' payoffs depend on the individual state variable this formulation includes the case in which there is a finite number of different types of small players.

Assumption 1 *v^b and v are continuous on M. Moreover* $0 \leq v, v^b \leq \bar{v}$.

A pure strategy for the large player is a mapping $\mathbf{y}_t : H_t \to Y$; a mixed (behavioral) strategy is a mapping $\boldsymbol{\sigma}_t : H_t \to \Sigma$. Similarly a strategy for a small player is a mapping[7] $\mathbf{x}_t : H_t \times Z \to X$. An aggregate strategy for the small players is a mapping $\boldsymbol{\mu}_t : H_t \to M$ that satisfies the following consistency requirement: For $h_t = (h_{t-1}; \mu_{t-1}, y_{t-1}, \lambda_t)$, we

[6] Notice that given the transition law λ_t is determined by $\mu_1, y_1, \ldots, \mu_{t-1}, y_{t-1}$.

[7] Given that private histories are unobservable, we assume that small players do not condition their play on their private history.

have $[\boldsymbol{\mu}_t(h_t)]^Z = \lambda_t$. In other words, for every history the marginal distribution of $\boldsymbol{\mu}_t(h_t)$ over states has to coincide with the current state λ_t. Finally, $\boldsymbol{\sigma} = (\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_t, \ldots)$, $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_t, \ldots)$, and $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_t, \ldots)$.

In an abuse of notation we will often write $V(\delta, \boldsymbol{\sigma}, \boldsymbol{\mu}, \mathbf{x}; h_t, z_t)$ as the expected payoff to a small player from playing $\mathbf{x}$, starting at state z_t after history h_t. Similarly $V^b(\beta, \boldsymbol{\sigma}, \boldsymbol{\mu}; h_t)$ is the expected payoff to the large player after history h_t.

2.1 Best Response and Aggregate Best Response

For a given strategy of the large player and a given aggregate strategy for the small players (which no individual small player can influence) we define an ϵ best response as follows:

Definition 1 (ϵ Best Responses) *The strategy $\mathbf{x} = (\mathbf{x}_t)_{t=1}^{\infty}$, is an ϵ best response for player i to $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ if for all $h_t \in H_t, t = 1, \ldots,$ such that h_t is a public history that is reached with strictly positive probability and for all $z \in Z$, $V(\delta, \boldsymbol{\sigma}, \boldsymbol{\mu}, \mathbf{x}; h_t, z) \geq V(\delta, \boldsymbol{\sigma}, \boldsymbol{\mu}, \mathbf{x}'; h_t, z) - \epsilon$, for all $\mathbf{x}'$. Let $B^{\epsilon}(\boldsymbol{\sigma}, \boldsymbol{\mu}; \lambda, z)$ denote the set of best responses given $\boldsymbol{\sigma}$, $\boldsymbol{\mu}$, and initial state λ, z.*

Let $B_t^{\epsilon}(\boldsymbol{\sigma}, \boldsymbol{\mu}; h_t, z)$ be the ϵ best response in period t only, i.e. :

$$B_t^{\epsilon}(\boldsymbol{\sigma}, \boldsymbol{\mu}; h_t, z) = \{x \in X \mid \mathbf{x}(h_t, z) = x, \text{ for some } \mathbf{x} \in B^{\epsilon}(\boldsymbol{\sigma}, \boldsymbol{\mu}; z)\}$$

Note that for $\epsilon = 0$ we have the conventional best response.

Definition 2 (Aggregate ϵ Best Response) *The aggregate strategy $\boldsymbol{\mu} = (\boldsymbol{\mu}_t)_{t=1}^{\infty}$, is an aggregate ϵ best response to $\boldsymbol{\sigma}$ for initial state λ, if for all h_t that are reached with strictly positive probability there is a μ with $|\mu - \boldsymbol{\mu}_t(h_t)| < \epsilon$ such that $x \in B_t^{\epsilon}(\boldsymbol{\mu}, \boldsymbol{\sigma}; h_t, z)$, for all $(x, z) \in$ suppμ. Let $E^{\epsilon}(\boldsymbol{\sigma}, \lambda)$ denote the set of aggregate ϵ best responses to $\boldsymbol{\sigma}$ for initial state λ.*

Therefore an aggregate ϵ best response to a strategy $\boldsymbol{\sigma}$ of the large player is an aggregate strategy $\boldsymbol{\mu}$ such that almost all individual strategies in its support are an ϵ best response to $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ for all reached histories.

Finally let $E_t^{\epsilon}(\boldsymbol{\sigma}, h_t)$ be defined as the aggregate ϵ best response in period t only, given a history h_t:

$$E_t^{\epsilon}(\boldsymbol{\sigma}, h_t) = \{\mu \in M | \boldsymbol{\mu}(h_t) = \mu \text{ for some } \boldsymbol{\mu} \in E^{\epsilon}(\boldsymbol{\sigma})\}$$

When $t = 1$, $h_1 = \{\lambda\}$, therefore we will write $E_1^{\epsilon}(\boldsymbol{\sigma}, \lambda)$ instead of $E_1^{\epsilon}(\boldsymbol{\sigma}, h_1)$.

For $\epsilon = 0$ we get the usual best response and aggregate best response. Let B, B_t, E, E_t, denote the best response and aggregate best response for $\epsilon = 0$.

Note that according to Definition 2, all small players may be able to gain ϵ *every period* if an aggregate ϵ best response is played. Thus for an aggregate ϵ best response to be close to an aggregate best response ϵ has to be small *relative to the discount factor* δ since $\epsilon/(1-\delta)$ measures the maximum utility loss for a typical small player over the course of the game. While for a fixed discount factor an ϵ can be chosen such that $\epsilon/(1-\delta)$ is very small, when we will consider the case where the discount factor of the small players is arbitrarily close to one (Section 7), we will need to use a stronger notion of aggregate ϵ best response.

3 The Perturbed Game

Now we consider a slight variation of the game defined above. Suppose that the small players are not completely sure about the large player's payoff function and in particular that they believe that with positive probability the large player's payoff function is different from the one described in the previous section. Let Ω be the space of potential types with generic element ω. Then the large player's payoff function will also depend on his type, $V^b(\beta, \mathbf{y}, \boldsymbol{\mu}; \omega, h_t)$. Let ω_0 denote the event that the type of the large layer is such that his payoff function is like in the unperturbed game, i.e. $V^b(\beta, \mathbf{y}, \boldsymbol{\mu}; \omega_0, h_t) = V^b(\beta, \mathbf{y}, \boldsymbol{\mu}, h_t)$. In the following we will call type ω_0 the rational or normal player.

Types other than ω_0 may have a possibly history dependent payoff function that makes a given pure strategy dominant in the infinite game. Such players will be called commitment players and for the sake of simplicity will be identified by the strategy they play rather than by their payoff function.

The existence of these commitment types captures uncertainty of the small players about the type of large player they are facing. The idea is that although the small players are almost certain they face the rational type, they cannot exclude the possibility that the large player perceives the game in a different way and hence will behave "irrationally". To account for the possibility that the large player can be of different types we have to modify the definition of a strategy for the large player. A mixed (behavioral) strategy for the large player is now a mapping $\boldsymbol{\sigma}_t : H_t \times \Omega \to \Sigma$.

Since the small players cannot observe the type of the large players, the definition of a strategy for the small players remains unchanged. We assume that the prior distribution of types is common knowledge. By $\sigma \backslash \sigma'(\omega)$ we denote the strategy that is obtained by substituting $\sigma'(\omega)$ for $\sigma(\omega)$ in σ.

Definition 3 *A Nash equilibrium for initial state λ is a (σ, μ) with $\mu_1^Z = \lambda$, such that $\mu \in E(\sigma, \lambda)$, and for all $\omega \in \Omega$, $V^b(\beta, \sigma, \mu; \omega, \lambda) \geq V^b(\beta, \sigma \backslash \sigma'(\omega), \mu; \omega, \lambda)$ for all $\sigma'(\omega)$.*

First we want to investigate the consequences of imitating a particular commitment type on the beliefs of the small players. In Lemma 1 we show that if the large player chooses to imitate a pure strategy of a particular commitment type, then in all but finitely many periods the small players will actually believe that with high probability the aggregate play will be consistent with this strategy being played in the next τ periods. Both the formulation and the proof of Lemma 1 are an extension of a result in Fudenberg and Levine (1989).

Let $\mathbf{y}^*$ denote the pure strategy played by a particular commitment type ω^*. Let h^* be the event that $y_t = \mathbf{y}_t^*(h_t)$ for all h_t that are reached following $(\mathbf{y}^*, \mu)$ starting from a given $h_0 = \lambda_0$. Furthermore let $p(\omega^*) = p^*$ denote the prior probability that $\omega = \omega^*$. Let $\pi_t^{*\tau}$ be the probability that in the next τ periods the actions of the large player are consistent with $\mathbf{y}^*$, i.e. the probability that in periods $t, t+1, \ldots, t+\tau-1$ aggregate play is consistent with $\mathbf{y}^*$ being played given the aggregate strategy μ, i.e. $\pi_t^{*\tau} = Pr[y_t = \mathbf{y}_t^*(h_t), \ldots, y_{t+\tau-1} = \mathbf{y}_{t+\tau-1}^*(h_{t+\tau-1}) | h_{t-1}, \mu]$. Finally, let $n(\pi_t^{*\tau} \leq \bar{\pi})$ be the random variable denoting the number of periods in which $\pi_t^{*\tau} \leq \bar{\pi}$.

Lemma 1 *Let $0 < \bar{\pi} < 1$ and suppose that $p^* > 0$, and that (σ, μ) are such that $Pr(h^* | \omega^*) = 1$. Then*

$$Pr\left[n(\pi_t^{*\tau} \leq \bar{\pi}) > \tau \frac{\log p^*}{\log \bar{\pi}} | h^*\right] = 0.$$

Remark: Note that since certain states may not be reached along a given history h^*, the small players will not get convinced that the large player actually uses the same strategy as the commitment type. However, since no individual small player can affect the aggregate state, the play in public histories that are not reached is irrelevant for any small player's decision problem.

4 The Case with No Strategic Externality

In this section we consider the simpler case in which the payoff of every small player is independent of the aggregate play of the small players.

Assumption 2 (No Strategic Externality) *v is independent of μ.*

The following equation defines $\bar{V}^b$ to be the limit of time averages of the large player's payoff. Since time averages need not converge we will take the limit infimum. Let

$$\bar{V}^b(\mathbf{y}, \boldsymbol{\mu}; \lambda) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} v^b(y_t, \mu_t)$$

where (y_t, μ_t) is the history induced by $\mathbf{y}$ and $\boldsymbol{\mu}$ and λ.

Define a strategy $\mathbf{y}$ such that $\mathbf{y}_t(h_t) = y_t$ for all h_t a *simple strategy.* A simple strategy is a strategy that does not depend on history but only on calendar time. With an abuse of notation in the following we will sometime identify a simple strategy with the infinite sequence of actions it prescribes, $\mathbf{y} = y$.

Let $\bar{V}^b(\epsilon, \lambda)$ be the best time average the large player could guarantee to himself by committing to a given simple strategy subject to the small players playing an aggregate ϵ best response. Then

$$\bar{V}^b(\epsilon, \lambda) = \sup_{y \in Y^\infty} \inf_{\mu \in E^\epsilon(y, \lambda)} \bar{V}^b(y, \mu; \lambda).$$

Let $\bar{V}^b(\epsilon) = \inf_\lambda \bar{V}^b(\epsilon, \lambda)$ and let $\bar{V}^b = \lim_{\epsilon \to 0} \bar{V}^b(\epsilon)$.

Now we define a collection of types (the Stackelberg types) which can be used by the rational large player to establish a reputation. Let $y(\epsilon, \eta, \lambda) = (y_1(\epsilon, \eta, \lambda), \ldots, y_t(\epsilon, \eta, \lambda), \ldots)$ be a simple strategy that satisfies

$$\inf_{\boldsymbol{\mu} \in E^\epsilon(y(\epsilon, \eta, \lambda), \lambda)} \bar{V}^b(y(\epsilon, \eta, \lambda), \boldsymbol{\mu}, \lambda) \geq \bar{V}^b(\epsilon, \lambda) - \eta.$$

Hence $y(\epsilon, \eta, \lambda)$ is an "almost" optimal sequence if the criterion is the limit of time averages.

The type $\omega(\epsilon, \eta, T)$ is defined by the following strategy:

- In the first T periods this type follows $y(\epsilon, \eta, \lambda_1)$.

- In case the small players reacted with an ϵ best response in period 1, this type continues with $y(\epsilon, \eta, \lambda_1)$ in period $T + 1$. If the small players did not choose an action close to a best response in period 1, this commitment type switches to $y(\epsilon, \eta, \lambda_{T+1})$.

- The same pattern is repeated for all periods: The commitment type will continue following the sequence $y(\epsilon,\eta,\lambda)$ if either it has been played for fewer than T periods or if T periods ago "almost" a best response was played. Otherwise a new sequence $y(\epsilon,\eta,\lambda')$ will be started, where λ' is the current value of the state variable.

More precisely, let $\theta_t \in \mathbf{N}$ $t = 1, 2, \ldots$ be defined as follows:

$$\theta_t = \begin{cases} 1 & \text{if } t = 1 \\ \theta_{t-1} + 1 & \text{if } \theta_{t-1} < T \text{ or if } \mu_{t-T} \in E^{\epsilon}_{t-T}(y(\epsilon,\eta,\lambda_{t-\theta_{t-1}}), h_{\tau-T}) \\ 1 & \text{otherwise} \end{cases}$$

Now let the type $\omega(\eta,\epsilon,T)$ be defined by following strategy:

$$\mathbf{y}_t(h_t) = y_{\theta_t}(\epsilon,\eta,\lambda_{t-\theta_t})$$

Note that the T-period lag in adjusting the optimal policy in the definition of the commitment types is crucial to avoid the time-inconsistency problem. A commitment type who chooses the optimal policy for the current state in every period is of little use to the large player since he wants to commit to ex-ante rather than ex-post optimal policies.

Assumption 3 *For all (ϵ,η) there is an $\epsilon' < \epsilon, \eta' < \eta$ such that $\omega(\epsilon',\eta',T) \in \Omega$ has strictly positive prior probability for all finite T.*

Assumption 3 says that there is a large variety of the described Stackelberg types. In particular, for arbitrarily small (ϵ,η) we can find a commitment type with positive prior for any finite "lag parameter" T.

In Theorem 1 we make two important assumptions that will be relaxed later. First we assume that v is independent of μ, i.e. there is no "strategic externality" in the play of the small players. Second we assume that the small players' discount factor stays fixed while the large player's discount factor approaches 1. Theorem 1 states that if the type space includes a particular collection of commitment types then as the discount factor of the large player goes to 1 in any Nash equilibrium he gets at least $\bar{V}^b$.

Theorem 1 *Suppose that Assumptions 1, 2 and 3 hold. Then in any Nash equilibrium $(\boldsymbol{\sigma},\boldsymbol{\mu})$ for initial state λ, $\lim_{\beta\to 1} V^b(\beta;\boldsymbol{\sigma},\boldsymbol{\mu},\lambda) \geq \bar{V}^b$.*

The intuition behind Theorem 1 is the following: Suppose that the large player imitates a commitment type $\omega(\epsilon,\eta,T)$. Then Lemma 1 implies that after a finite number of periods the small players will actually believe that the large player will continue to play like the commitment type in the next $\hat{T}$ periods with very high probability. Since the small players discount future payoffs there is a $\hat{T}$ large enough such that if the commitment strategy is followed with large probability in the next $\hat{T}$ periods then the small players will actually play a best response to the commitment strategy. Since we can find commitment types with positive prior probability for arbitrarily large lag parameter T, we can choose T in such a way that $T > \hat{T}$. In this case an aggregate ϵ best response to the commitment strategy implies an aggregate best response to an optimal sequence $y(\epsilon, \eta, \lambda)$. The Theorem then follows from the fact that this argument can be repeated for arbitrarily small (ϵ, η).

5 Firm Specific Human Capital Accumulation: An Example

A large amount of the accumulation of human capital takes place "on the job". Becker (1975) distinguishes between general formation and firm specific formation and argues that general formation is paid for by workers (in the form of a wage lower than the opportunity wage) while specific formation is paid for by the firm. The accumulation of firm specific human capital is finally argued to be an explanation of the relationship between job qualifications and turn over rate and of labor hoarding in the business cycle.

An important assumption of this analysis is that the human capital investment level can actually be observed by the firm. However, since any human capital investment requires at least some degree of cooperation on the part of the employee, it is probably more reasonable to assume that the firm cannot observe the employee's investment effort but only his productivity level.

The purpose of this section is to apply the results of section 4 to show that even when a firm can only observe the productivity levels of its employees, Becker's (1975) result is still true in the sense that firm specific capital accumulation does take place and that an arbitrarily patient firm appropriates all its net value.

Consider the following strategic situation: A firm is associated with a large number of identical workers; each worker can invest in improving his skills or his product. This investment will typically be productive only after one or more periods. The firm can observe the productivity of each worker but it cannot observe the investment he makes in improving his product. The firm has to decide a wage schedule each period, i.e. it associates to each productivity level a wage.

Let b denote the firm and let $i \in [0,1]$ denote the workers. Each worker in each period is in one of $H+1$ different states (productivity levels). Let $Z = \{z^0, \ldots, z^h, \ldots, z^H\}$ denote the set of productivity levels. The worker can choose among $J+1$ different investment levels each period. Let $X = \{x^0, \ldots, x^J\}$ denote the set of possible investment levels of the workers. Let $x^0 = 0$. A worker in state $z \in Z$ produces an output of size z in the current period. Let z_t be the state of a worker in period t. His state in period $t+1$ is given by $z_{t+1} = f(z_t, x_t)$.

The firm cannot observe individual histories. It can, however, observe the distribution over productivity levels in every period. At the beginning of each period the firm observes the distribution λ over productivity levels. The firm chooses a wage schedule that associates a wage with every productivity level. Let $W = \{w^0, \ldots, w^N\}$ be the set of possible wages. We assume that the grid on W is "fine", i.e. $w^n - w^{n-1} < \xi$ and that $x^j \in W, \forall j = 0, \ldots, J$. The set of actions for player b is given by W^{H+1}, i.e. $Y = W^{H+1}$. For $y \in Y$, let $y(z)$ be the wage schedule corresponding to state z. For a worker who chooses x_t in period t and is in state z_t the payoff in period t is given by[8]

$$v(y_t, x_t, z_t) = y(z_t) - x_t$$

Given the aggregate state is λ_t and the firm chooses y_t the payoff to b in period t is given by:

$$v^b(y_t, \lambda_t) = \sum_Z \lambda_t(z)(z - y_t(z))$$

Note that the workers' payoff is independent of the aggregate state.

Example 1 The following example illustrates how the subgame perfect equilibria in this dynamic game may not contain any efficient equilibrium.

Let $X = \{0, x\}, Z = \{z^l, z^h\}, f(z^l, 0) = z^l, f(z^H, 0) = z^H, f(z, x) = z^H, \forall z \in Z$. Let $\lambda = \lambda(z^h)$ be the proportion of high productivity workers.

Proposition 1 *$\lambda_t = \lambda_1$ and $y_t = 0$ is the unique subgame perfect equilibrium in this game, for all $\delta, \beta \in [0,1)$.*

Proof: Note that for every history $h \in H$, λ_t is an increasing sequence. Thus $\lambda_t \to \hat{\lambda}$ for some $\hat{\lambda}$. If $\hat{\lambda} = \lambda_1$ there is nothing to prove. Hence assume $\hat{\lambda} > \lambda_1$. If $\lambda_t \to \hat{\lambda}$, for every

[8] Since v does not depend on the aggregate action, Assumption 2 is satisfied.

$\epsilon > 0$ there is a t such that $\hat{\lambda} - \lambda_t < \epsilon$ for all $\tau > t$. Choose an $\epsilon > 0$ such that

$$\epsilon z^H < (1 - \beta)(\hat{\lambda} - \epsilon)x$$

Now let t be such that $\hat{\lambda} - \lambda_t < \epsilon$. Since the left hand side of the above inquality gives an upper bound on what the firm could gain by setting a $y > 0$ and the right hand side gives a lower bound on the cost to the firm of setting a y such that the workers will be willing to invest, in period t only $y_t = 0$ can be optimal. Thus in period $t-1$, $x_{t-1} = 0$ and $\lambda_t = \lambda_{t-1}$. Hence by induction $\hat{\lambda} = \lambda_1$. □.

Even though for

$$\frac{\delta z^h}{1 - \delta} > x$$

moving to state z^h would be efficient, there is no subgame perfect equilibrium that can support this move.

5.1 Using Reputation

In the following we show how in this simple class of games reputation can be used to find a tight bound on equilibrium payoffs.

Let $(z_1(\beta), (x_t(\beta))_{t=1}^{\infty})$ be the pair such that for $z_{t+1} = f(z_t, x_t)$

$$\sum_{t=1}^{\infty} \beta^{t-1}(z_t(\beta) - x_t(\beta)) \geq \sum_{t=1}^{\infty} \beta^{t-1}(z_t - x_t)$$

for all $(z_1, (x_t))$. Hence $(z_1(\beta), (x_t(\beta)))$ is the state investment pair that achieves the largest net surplus (discounted with the firm's discount factor).

Assumption 4 *There is a $\hat{\beta} < 1$ such that for $\beta \in [\hat{\beta}, 1)$, $z_1(\beta) = z^*, x_t(\beta) = x^*$, for all t and $f(z^*, x^*) = z^*$. Furthermore for all $z \in Z$ there is a feasible sequence $(z_t, x_t)_{t=1}^{\bar{T}}$ with $\bar{T} \leq H + 1$ such that $z_0 = z$ and $z_{\bar{T}} = z^*$.*

The first part of the assumption says that the optimal state/investment pair is stationary, i.e. there is a (z^*, x^*) such that along the optimal sequence (z^*, x^*) will be repeated every period. Note, however, that results of the same kind of the ones we will discuss in Proposition 2 can be easily obtained at the expense of a heavier notation without the first part of Assumption 4. The second part of Assumption 4 says that it is possible to reach state z^* from every other state z.

Let $c(\delta, z)$ be the cost of moving from state z to z^*, i.e.

$$c(\delta, z) = \min_{(x_1,\ldots,x_k)} (1-\delta)\sum_{j=1}^{k} \delta^{j-1} x_j$$

$$\text{s.t. } z_1 = z,\ z_j = f(z_{j-1},\ x_{j-1}), z_k = z^*.$$

We assume that $c(\delta, z) > (1-\delta)x^*$ for all $z < z^*$, i.e. it is less costly to stay at the productivity level z^* than to move to z^* from a less productive state.

Let $\bar{\delta}$ satisfy the following equations:

$$\max_z c(\bar{\delta}, z) \quad < \quad \bar{\delta}^{H+1}\xi \tag{1}$$

$$\bar{\delta} \min_{z<z^*} c(\bar{\delta}, z) \quad > \quad (1-\bar{\delta})x^* \tag{2}$$

Since $c(\delta, z) \to 0$ as $\delta \to 1$ there exists a $\bar{\delta} < 1$ such that inequalities (1) and (2) are satisfied for all $\delta > \bar{\delta}$.

Suppose the type ω^* (the "Stackelberg" type) follows the following strategy:

$$y_t(z) = \begin{cases} x^* + \xi & \text{if } z = z^* \\ w^0 & \text{otherwise} \end{cases}$$

The strategy $\mathbf{y}$ imposes a simple two-tier system: unexperienced workers get the reservation wage and experienced workers get their investment expenses plus a small premium that is just sufficient to give them an incentive to accumulate (if $\delta > \bar{\delta}$). Furthermore if $\mathbf{y}$ is played and $\delta > \bar{\delta}$ then it is optimal for every worker to stay in state z^*, i.e. to invest x^*, since inequality (2) holds.

The following proposition gives precise bounds for the payoff of the firm as the discount factor of the firm approaches 1.

Proposition 2 *Suppose Assumption 4 holds and suppose that $\delta > \bar{\delta}$. Let V^b denote the expected payoff of the rational firm in period 1. If ω^* has prior probability $p^* > 0$, then in any Nash equilibrium $z^* - x^* \geq \lim_{\beta\to 1} V^b \geq z^* - x^* - \xi$.*

Proof: The first inequality follows from the fact that the expected payoff of any worker has to be nonnegative since every worker has the option of choosing $x_t = 0$, for all t and therefore the inequality follows from Assumption 4. The second inequality follows from Theorem 1. □

Proposition 2 shows that if there is an arbitrarily small probability that the firm is a type that plays an appropriate strategy, the firm can actually improve upon all the subgame perfect equilibrium payoffs in the unperturbed game by establishing a reputation for playing like the commitment firm.

6 Including Strategic Externalities

In many economic problems the utility of individuals depends on their individual choice as well as on the aggregate behavior of the other individuals, for example through prices in a market game or through the aggregate level of capital in a capital accumulation problem with externalities. In this section we discuss reputational arguments in the general case in which the payoffs to individual small players may also depend on the aggregate play of the small players.

Allowing for this possiblity complicates the analysis for the following reason: Even though small players discount future payoffs at a fixed rate $\delta < 1$, it is not true that the aggregate ϵ best response *today* does not change when the large player's action in the very distant future changes.

The large player can only exploit his reputation successfully if the small players choose an aggregate ϵ best response to the commitment strategy whenever the large player imitates this strategy long enough. This implies that we need to find a (uniform) bound T such that if the small players believe that the commitment strategy is played for the next T periods, then they will actually play an aggregate ϵ best response to it. When v is independent of μ discounting implies that we can find such a T uniformly over all strategies. If v depends on μ, this property fails. The following example illustrates this point.

Example 2 Consider an economy in which there is a continuum of private agents (small players) and a government (large player). Suppose that private agents can choose between becoming *specialized* or staying *autarkic*. Then the strategy space for each private agent is $X = \{0, 1\}$ where $x = 0$ symbolizes autarky and $x = 1$ specialization. Any agent can either be in an experienced state, $z = 1$, or in an inexperienced state, $z = 0$. Experience is obtained after having specialized for one period. If an experienced agent fails to specialize, then he loses his experience. Hence the individual state variable transition can be summarized by: $f(x, z) = x$, $x = 0, 1$.

Only experienced agents who specialize are productive. However, their payoff from specialization depends on how many other agents decide to specialize in the current period (irrespective of whether these agents are experienced or not). Let $\mu^X(1)$ be the fraction of agents who specialize in the current period, then the value of the output produced by an agent who plays x and is in state z is: $\mu^X(1)zx - cx$ where c is the cost of specialization.

The government has two policies: It can either do nothing ($y = 0$) or it can reward all the experienced specializing agents by giving them a subsidy of 1 for each unit they produce ($y = 1$). With this set-up the payoff function of a private agent will be

$$v(\mu, y, x, z) = (\mu^X(1) + y)zx - cx.$$

The government is benevolent but giving a subsidy is costly. Let $\mu(1,1)$ denote the proportion of experienced private agents (private agents in state $z = 1$) who decide to specialize ($x = 1$). Then the government's payoff function can be written as

$$v^b(y, \mu) = \mu(1,1)(1 + (1-k)y) - c\mu^X(1)$$

where $k > 1$ is the unit cost of raising funds to pay the subsidies.

The following table summarizes the payoffs to the private agents. The column entries are combinations of actions and individual values of the state variable of the small player (z, x); the row entries denote actions of the large player.

	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$y = 0$	0	$-c$	0	$\mu_t^X(1) - c$
$y = 1$	0	$-c$	0	$\mu_t^X(1) + 1 - c$

Let $c < \delta < 1$. Under policy $y = 0$ the private agents will specialize ($x = 1$) only if enough other small players specialize. Under policy $y = 1$ there is a reward for experienced agents who specialize.

The government would like to play policy $y = 0$ in every period and would like all small players to specialize (choose action 1). However this is not the only aggregate best response to $\mathbf{y} = (0, 0, \ldots, 0, \ldots)$. Any sequence of the form:

$$\mu_t^X(1) = \begin{cases} 1 & \text{if } t \leq \tau \\ 0 & \text{if } t > \tau \end{cases}$$

for $\tau \geq 0$ is an aggregate best response. In particular $\mu_t^X(1) = 0$ for all t is the worst aggregate best response.

Now suppose the government plays $\mathbf{y} = (0, \ldots, 0, 1; 0, \ldots, 0, 1; \ldots)$, where the sequence of consecutive 0's is arbitrarily large. Whenever the government gives a subsidy ($y = 1$

is played) in period τ every small player wants to specialize ($x_\tau = 1$) *and be experienced* ($z_\tau = 1$). But this implies that in period $\tau - 1$ every small player has to specialize ($x_{\tau-1} = 1$), otherwise he would not be experienced in the following period. This in turn implies that also in period $\tau - 1$ every small player can benefit from specialization ($x_{\tau-1} = 1$) as long as he is experienced ($z_{\tau-1} = 1$). But to be experienced in period $\tau - 1$ ($z_{\tau-1} = 1$) he has to specialize in $\tau - 2$ ($x_{\tau-2} = 1$) and so on.

Thus every small player will choose $x_t = 1$ for $t \leq \tau$, which implies that the unique equilibrium is $\mu_t^X = 1$, for all t. In order to guarantee that the private agents will actually specialize ($x = 1$) the large player has to give a subsidy (switch to policy $y = 1$) every once in a while. •

Theorem 1 relied on the fact that we could find a uniform bound T such that the large player's actions more than T periods from now did not affect the small players' current behavior. The previous example shows that in the case with strategic externalities such a uniform bound does not exist[9]. This creates a problem for a large player who tries to exploit his reputations: the small players may have to be convinced that the large player follows a given strategy for very many future periods.

As an illustration, consider again Example 2. Suppose that the large player wants to establish a reputation for playing the sequence $y = (A, B; A, A, B; A, A, A, B; \ldots)$. Clearly $\mu_t^X(1) = 1$ is the unique aggregate ϵ best response to y. However, to ensure that this best response is played in period t, π_t^{y,T_t}, with $T_t \to \infty$ has to be sufficiently large. If T_t goes to infinity very fast then the large player may actually never be able to establish a sufficiently "far-reaching" reputation so that the small players will play $\mu_t^X(1) = 1$.

To circumvent this problem we will assume that by committing to an "eventually periodic" sequence the large player can do almost as well as by committing to an arbitrary sequence. This allows us to restrict the Stackelberg type to a set of strategies for which we can find a uniform bound on the number of future periods that matter for the current behavior of the small players.

Recall that a pure strategy $\mathbf{y}$ for b is called simple, if $\mathbf{y} \equiv y$, for some $y \in Y^\infty$; i.e. no matter what history is reached in period t, player b chooses y_t in period t. Define a simple strategy L *periodic* if for some $l, k \leq L$, $L < \infty$, we have $y_{t+l} = y_t$ for all $t \geq k$. Let $Y(L)$ denote the set of L *periodic simple strategies.* The following assumption says that by

[9]In other words: the aggregate ϵ best response fails to be lower hemi continuous in the product topology.

committing to an L periodic sequence, the large player can guarantee himself almost the same payoff as by committing to an arbitrary sequence.

Assumption 5 *For all $\eta > 0, \epsilon > 0$, there is an L such that for all y, λ there is a $y' \in Y(L)$ such that $\inf_{\boldsymbol{\mu} \in E^\epsilon(y',\lambda)} \bar{V}^b(y', \boldsymbol{\mu}, \lambda) \geq \inf_{\boldsymbol{\mu} \in E^\epsilon(y,\lambda)} \bar{V}^b(y, \boldsymbol{\mu}, \lambda) - \eta$.*

Note that Assumption 5 is satisfied in Example 2.

Commitment types (Stackelberg types) are constructed analogous to the ones in Section 4. The only difference is that we restrict the Stackelberg type to the use of L periodic sequences. Let $y(\epsilon, \eta, \lambda) \in Y(L)$ satisfy

$$\inf_{\boldsymbol{\mu} \in E^\epsilon(y(\epsilon,\eta,\lambda),\lambda)} \bar{V}^b(y(\epsilon, \eta, \lambda), \boldsymbol{\mu}, \lambda) \geq \bar{V}^b(\epsilon, \lambda) - \eta$$

Assumption 5 guarantees the existence of such a sequence. As before we define $\theta_t \in \mathbf{N}$ $t = 1, 2, \ldots$ as follows:

$$\theta_t = \begin{cases} 1 & \text{if } t = 1 \\ \theta_{t-1} + 1 & \text{if } \theta_{t-1} < T \text{ or if } \mu_{t-T} \in E^\epsilon_{t-T}(y(\epsilon, \eta, \lambda_{t-\theta_{t-1}}), h_{\tau-T}) \\ 1 & \text{otherwise} \end{cases}$$

Type $\omega(\epsilon, \eta, T)$ is committed to the strategy:

$$\mathbf{y}_t(h_t) = y_{\theta_t}(\epsilon, \eta, \lambda_{t-\theta_t})$$

The interpretation of this strategy is the same as the one that was provided for the case with no strategic externality.

Theorem 2 *Suppose that Assumptions 1,3, and 5 hold. In any Nash equilibrium $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ for initial state λ, $\lim_{\beta \to 1} V^b(\beta; \boldsymbol{\sigma}, \boldsymbol{\mu}, \lambda) \geq \bar{V}^b$.*

Theorem 2 generalizes Theorem 1 to include the possibility of the small player's payoff to depend on the aggregate play of the small players. If Assumption 5 holds, then Theorem 2 says that as the discount factor of the large player goes to 1 in any equilibrium he gets at least the time average of payoffs corresponding to an optimal commitment.

7 Patient Small Players

In Theorem 1 we assumed that the discount factor of the small players stays fixed while the large player becomes arbitrarily patient. In applications like policy games the utility function of the large player frequently reflects the utility function of the small players (e.g.

the large player's preferences are identical to the utility function of the "median voter"). Thus it is important to identify classes of games where reputation allows the large player to achieve essentially his commitment payoff when both the large and the small players become arbitrarily patient simultaneously.

The difficulty in establishing a reputation with patient small players lies in the fact that small players may become increasingly reluctant to take an action that leads to an irreversible state as they get more patient. Thus to convince a very patient small player to take this action the large player may have to establish a reputation for following the commitment strategy for very many periods and hence it may take "too long" to establish a reputation that induces the small players to enter an irreversible state.

7.1 The Failure of Reputation in the Durable Goods Monopoly

The following simple example of a monopolist selling a durable good to a population of buyers illustrates the failure of reputational arguments with irreversible states.

Example 3 Suppose there are two types of buyers H and L. The reservation price of type H, r^H, for the durable good is 5, the reservation price of type L, r^L, is 2. There is mass 1/2 of both types of buyers. Each period the buyer takes either action 0 (he does not buy) or action 1 (he buys). Similarly the state of a buyer is either 0 (no purchase has occured in the past) or 1 (a purchase has occured in some previous period). Thus the transition function is defined as:

$$f(x_t, z_t) = \begin{cases} 0 & \text{if } x_t = 0 \text{ and } z_t = 0 \\ 1 & \text{otherwise} \end{cases}$$

The monopolist sets a price p_t every period, where $p_t \in \{0, 1/n, \ldots, (5n-1)/n\}; n \geq 6$. If buyer $i \in \{H, L\}$ purchases the durable good in period t then his payoff is $\delta^t(r^i - p_t)$. More precisely, buyer i's payoff function is

$$v^i(p_t, x_t, z_t) = \begin{cases} r^i - p_t & \text{if } z_t = 0 \text{ and } x_t = 1 \\ -p_t & \text{if } z_t = 1 \text{ and } x_t = 1 \\ 0 & \text{otherwise} \end{cases}$$

For any sequence of prices p and aggregate actions μ the payoff to the large player is[10]

$$V^b(p, \mu) = \sum_{t=1}^{\infty} \delta^{t-1} \cdot p_t \cdot \mu_t(1, 0)$$

[10] We do not include $\mu_t(1,1)$ (the proportion of buyers that have already bought the durable good in the past that do so again) because no buyer will purchase twice in equilibrium.

Suppose there are three types of monopolists: one normal type characterized by the payoff function above, type ω^* who sets $p_t = (5n-1)/n$ for all t and type $\hat{\omega}$ who follows the strategy:

$$p_t = \begin{cases} (5n-1)/n & \text{if } t \leq T \\ (2n-1)/n & \text{otherwise} \end{cases}$$

where $\log(1/2)/\log(\delta) < T < \log(2/(3n+1)))/\log\delta$. Both commitment types have prior probability $\epsilon > 0$.

The strategy of playing $p_t = (2n-1)/n$ (for the normal type) constitutes a sequential equilibrium for large δ. To see this first note that $p_t = (2n-1)/n$ constitutes a subgame perfect Nash equilibrium in the game where there is only the normal type if δ is sufficiently large. Thus it remains to show that the normal type does not have an incentive to imitate type ω^*. Suppose b deviates and offers $p_t = (5n-1)/n$. Since

$$\frac{\epsilon}{2 \cdot \epsilon} \cdot \delta^T \cdot (5 - (2n-1)/n) > 1/n$$

type H will not buy until period $T + 1$[11]. However

$$\delta^T \frac{1}{2} \cdot \frac{5n-1}{n} + \delta^{T+1} \frac{1}{2} \cdot \frac{2n-1}{n} < \frac{7n-1}{4n} < \frac{2n-1}{n}$$

for $n \geq 6$, where the first element of the chain of inequalities is an upper bound on the payoff to b from deviating and the last is the payoff from setting $p_1 = (2n-1)/n$. This implies that deviation from $p_1 = (2n-1)/n$ does not pay. Thus in this game the large player is unable to exploit reputational effects to achieve the simple monopoly payoff $(5n-1)/2n$.$\square$

7.2 No Irreversible Actions

The following Assumption says that no action that the small players can take has irreversible consequences.

Assumption 6 (Reversibility of Accumulation Paths) *Suppose there is a sequence* $(y_1, \ldots, y_N)$, $(x_1, \ldots, x_N)$ *such that for* $z_1 = z$ *and* $z_n = f(y_n, x_n, z_{n-1})$ *we have* $z_N = z'$. *Then for any other sequence* $(\hat{y}_1, \hat{y}_2, \ldots)$ *there is a sequence* $(\hat{x}_1, .., \hat{x}_{N'})$ *such that for* $z_1 = z'$ *and* $z_n = f(\hat{y}_n, \hat{x}_n, z_{n-1})$ *we have* $z_{N'} = z$.

[11] The left hand side of the inequality is a lower bound on the expected payoff from waiting until $T+1$ and then buying at $p_{T+1} = (2n-1)/n$, and the right hand side is the payoff from buying at $(5n-1)/n$ at $t = 1$.

Using Assumption 6 we can partition the states Z into subsets Z^j such that the small players can only move between states in the same subset Z^j and furthermore there is an N such that for any pair (z, z') belonging to the same Z^j a small player can move from z to z' in fewer than N periods (independent of y).

Note that the definition of aggregate ϵ best response (Definition 2) contains strategies in which every small player "loses" ϵ units of utility as compared to a best response *every period.* Since δ is fixed in Theorem 1, we can make $\epsilon/(1-\delta)$ arbitrarily small. (Note that $\epsilon/(1-\delta)$ denotes the overall "loss" of utility of a typical small player as compared to a best reponse). Here we want to let $\delta \to 1$ and therefore we need a stronger notion of aggregate ϵ best response.

Denote by $y^T = (y_1, \ldots, y_T)$ a T period sequence of actions for the large player and similarly $\mu^T = (\mu_1, \ldots, \mu_T), x^T = (x_1, \ldots, x_T), z^T = (z_1, \ldots, z_T)$. Finally let $G^T(y^T) = \{(x^T, z^T) : z_{t+1}^T = f(y_t^T, x_t^T, z_t^T)\}$ denote the set of sequences (x^T, z^T) of length T that are feasible under y^T. Now define a truncated aggregate ϵ best response in the following way:

Definition 4 (Truncated Aggregate ϵ Best Responses) *μ^T is a T-truncated aggregate ϵ best response to y^T for initial state λ if $\mu_{t+1}^{TZ} = F(y_t^T, \mu_t^T), t = 1, \ldots, T-1$ and for all $(x^T, z^T) \in \text{supp}\mu^T \cap G^T(y^T)$*

$$\frac{1}{T}\sum_{t=1}^{T} v(y_t^T, x_t^T, z_t^T) \geq \frac{1}{T}\sum_{t=1}^{T} v(y_t^T, x_t'^T, z_t'^T) - \epsilon$$

for all $(x'^T, z'^T) \in G(y^T)$ with $z_1'^T = z_1^T$. Let $E^{T,\epsilon}(y^T, \lambda)$ denote the set of T-truncated aggregate ϵ best responses to y^T for initial state λ.

This definition of a truncated aggregate ϵ best response says that over the course of T periods the *average* payoff could not be increased by more than ϵ by any other sequence of actions. Note that this definition requires ϵ optimality (in a time average sense) over T periods and irrespective of the continuation of play and hence is a much stronger notion of aggregate ϵ best response than the one used in Theorems 1 and 2.

Next we define the limit of the commitment payoffs of a sequence of truncated games. Consider a truncated game in which the large player commits to an optimal sequence and the small players choose a truncated aggregate ϵ best response. $\hat{V}^b$ denotes the limit of payoffs for the large player when the game is truncated farther and farther away in the future. Again, since time averages need not converge, we take the limit infimum. Let

$$\hat{V}^b(\epsilon, \lambda) = \liminf_{T\to\infty}\{\max_{y^T} \min_{\mu^T \in E^{T,\epsilon}(y^T,\lambda)} \frac{1}{T}\sum_{t=1}^{T} v^b(y_t^T, \mu_t^T)\}$$

and

$$\hat{V}^b(\epsilon) = \inf_{\lambda} V^b(\epsilon, \lambda).$$

$\hat{V}^b = \lim_{\epsilon\to 0} \hat{V}(\epsilon)$.

Again we define a collection of commitment types who will allow the large player to establish a reputation. Let $y^T(\epsilon, \lambda)$ be a T period sequence that solves

$$\max_{y^T} \min_{\mu^T \in E^{T,\epsilon}(y^T,\lambda)} \frac{1}{T}\sum_{t=1}^{T} v^b(y_t^T, \mu_t^T)$$

We define the type $\omega(\epsilon, T)$, to play the strategy

$$\mathbf{y}(\epsilon, T) = (y_1^T(\epsilon, \lambda_1), \ldots, y_T^T(\epsilon, \lambda_1); y_1^T(\epsilon, \lambda_{T+1}), \ldots, y_T^T(\epsilon, \lambda_{T+1}); \ldots)$$

The commitment type $\omega(\epsilon, T)$ plays the optimal sequence in the T-period truncated game given the initial state at the beginning of the truncated game.

Assumption 7 *For all $\epsilon > 0$ there is an $\epsilon' < \epsilon$ such that $\omega(\epsilon', T) \in \Omega$ has strictly postive prior probability for every finite T.*

Theorem 3 says that if both players are very patient and if the transition function is reversible then in any Nash equilibrium the large player will receive at least a payoff that is close to the maximal time average in the T-period truncated game for arbitrarily large T.

Theorem 3 *Suppose Assumptions 1, 2, 6, and 7 hold and all players have a common discount factor δ. Then in any Nash equilibrium $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ for initial state λ, $\lim_{\delta\to 1} V^b(\delta; \boldsymbol{\sigma}, \boldsymbol{\mu}, \lambda) \geq \hat{V}^b$.*

The idea behind the proof of Theorem 3 is that we split up the infinite game into finite "superstage games" of length T. Note that the effect of a current decision of a small player on the payoffs in future "superstage" games can be "undone" in N periods or less by the reversibility assumption. If N is small as compared to T then the small players will behave almost like short lived players in every superstage game, i.e. they will behave essentially as if they were alive only for one superstage game. Therefore the large player can exploit his reputation if he convinces the small players that he will follow the commitment strategy in the current superstage game. Thus it is sufficient for the large player to establish a reputation for a bounded number of future periods and Lemma 1 shows that this can be accomplished in finitely many periods.

8 Conclusions

Whenever current play affects future payoff opportunities, agents' current decisions depend not only on present but also on future expected behavior of their opponents.

To describe this situation an infinite dynamic game between a large player and a continuum of small players has been studied and it has been shown that the use of reputational arguments allows to characterize the set of equilibria by providing a lower bound on the equilibrium payoffs to the large player. This has been accomplished by noticing that, if there is uncertainty relative to the type of the large player, the large player can actually establish a reputation for behaving in a certain way in a finite horizon.

An example has been presented to show that when individual small players' payoffs also depend on the aggregate play of the small players, it is possible that arbitrarily distant play of the large player affects current aggregate behavior of the small players. Interestingly it turns out that even in cases like this reputational arguments do have a bite: it is argued that the large player can establish a reputation for playing repeatedly an appropriate finite sequence of actions which in turn allows him to get at least his commitment payoff.

Provided that the small players' actions do not have irreversible consequences, reputational arguments have been shown to work independently of the rate of patience of the small players. Even when the large player and the small players have the same discount factor (like in the case of a benevolent government), the fact that the large player can establish a reputation for playing a strategy that depends on the aggregate state variable only after a sufficiently long adjustment lag provides a lower bound on the large player's equilibrium payoffs.

A simple example of a durable goods monopoly problem has been presented to illustrate the role of the assumption that the small players' actions do not have irreversible consequences: when some action profile leads to an absorbing state (purchase of the durable good in this case), then, if the small players are arbitrarily patient, the possibility of establishing a reputation may fail to improve the large player's payoff since no finite adjustment lag in the strategy of the large player would convince an arbitrarily patient small player to play a best response to the optimal commitment strategy. Only in a case like this is it possible that payoffs that are not close to the large player's optimal commitment payoff be equilibrium payoffs of the perturbed game.

9 Proofs:

9.1 Proof of Lemma 1:

Lemma 1 *Let* $0 < \bar{\pi} < 1$ *and suppose that* $p^* > 0$, *and that* $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ *are such that* $Pr(h^*|\omega^*) = 1$. *Then*

$$Pr\left[n(\pi_t^{*\tau} \leq \bar{\pi}) > \tau\frac{\log p^*}{\log \bar{\pi}}|h^*\right] = 0$$

Proof: Let $\bar{\omega}^*$ denote the event that $\omega \neq \omega^*$. Then by Bayes's law we have

$$\begin{aligned} Pr(\omega^*|h_{t+1}) &= Pr(\omega^*|\mathbf{y}_t^*(h_t), h_t) \\ &= \frac{Pr(\omega^*|h_t)Pr(\mathbf{y}_t^*(h_t)|\omega^*)}{Pr(\omega^*|h_t)Pr(\mathbf{y}_t^*(h_t)|\omega^*) + (1 - Pr(\omega^*|h_t))Pr(\mathbf{y}_t^*(h_t)|\bar{\omega}^*)} \end{aligned} \tag{3}$$

Notice that $Pr(\mathbf{y}_t^*(h_t)|\omega^*) = 1$ and that the denominator of (3) is equal to $Pr(\mathbf{y}_t^*(h_t))$. Therefore (3) can be rewriten as

$$Pr(\omega^*|h_{t+1}) = \frac{Pr(\omega^*|h_t)}{Pr(\mathbf{y}_t^*(h_t))}. \tag{4}$$

Notice that for any τ, $Pr(y_t = \mathbf{y}_t^*(h_t)) = Pr(y_{t'} = \mathbf{y}_{t'}^*(h_{t'}), t' = t, \ldots, t + \tau - 1) + Pr(\mathbf{y}_t^*(h_t), y_{t'} \neq \mathbf{y}_{t'}^*(h_{t'})$, for some $t' = t + 1, \ldots, t + \tau - 1)$. Recall $\pi_t^{*\tau} = Pr(y_{t'} = \mathbf{y}_{t'}^*(h_{t'}), t' = t, \ldots, t + \tau - 1)$, and let $\bar{\pi}_t^{*\tau} = Pr(y_t = \mathbf{y}_t^*(h_t), y_{t'} \neq \mathbf{y}_{t'}^*(h_{t'})$, for some $t' = t + 1, \ldots, t + \tau - 1)$, i.e. $\bar{\pi}_t^{*\tau}$ is the probability that the large player's play is in accordance with $\mathbf{y}^*$ at time t, but differ at some point in the next $\tau - 1$ periods.Then, for any fixed τ (4) can be rewritten as

$$Pr(\omega^*|h_{t+1}) = \frac{Pr(\omega^*|h_t)}{\pi_t^{*\tau} + \bar{\pi}_t^{*\tau}}.$$

Suppose that $\pi_{t'}^{*\tau} \leq \bar{\pi}$ for all $t' = t, \ldots, t + \tau - 1$. Then if the large player plays like the commitment type for $t' = t, \ldots, t + \tau - 1$ (i.e. $y_{t'} = \mathbf{y}_{t'}^*(h_{t'})$), then the probability that he is type ω^* has to go up by a factor of at least $1/\bar{\pi}$ (because if $y_{t'} = \mathbf{y}_{t'}^*(h_{t'})$ all $t' = t, \ldots, t + \tau - 1$, then at some $t' = t, \ldots, t + \tau - 1$ $\bar{\pi}_t^{*\tau}$ will be updated to zero. Given that $Pr(\omega^*|h_1) = p^*$, after τ periods

$$Pr(\omega^*|h_{\tau+1}) \geq p^*/\bar{\pi}.$$

If $\pi_t^{*\tau} \leq \bar{\pi}$ for τK periods during which $y_t = \mathbf{y}^*(h_t)$, all t, then

$$Pr(\omega^*|h_{K\cdot\tau+1}) \geq p^*/\bar{\pi}^K.$$

However, since

$$Pr(\omega^*|h_t) \leq 1 \tag{5}$$

if

$$p^*/\bar{\pi}^K > 1 \tag{6}$$

inequality (5) is violated and a contradiction to the hypothesis that $\pi_{t'}^{*\tau} \leq \bar{\pi}$ for all $t' = t, \ldots, t + K \cdot \tau - 1$ is obtained.

Taking the log of (6) the condition becomes

$$K > \log p^* / \log \bar{\pi}$$

and the proof is complete. □

9.2 Proof of Theorem 1

Theorem 1 *Suppose that Assumptions 1, 2 and 3 hold. Then in any Nash equilibrium* $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ *for initial state* λ, $\lim_{\beta \to 1} V^b(\beta; \boldsymbol{\sigma}, \boldsymbol{\mu}, \lambda) \geq \bar{V}^b$.

The strategy of the proof will be to show that if the large player imitates the Stackelberg type $\omega(\eta, \epsilon, T)$ for appropriately chosen T, then eventually the small players will play a best response to the Stackelberg strategy.

In the following we present a Lemma that shows that if the small players believe that the large player follows a given sequence of actions for a sufficient number of periods with a sufficiently large probability, then the small players will play an aggregate ϵ best response to this sequence of actions.

For a pure strategy $\mathbf{y}$, let $\pi_t^{\mathbf{y}T}$ be the probability that $\mathbf{y}$ is played in each of the next T periods, i.e. in the periods $t, t+1, t+2, \ldots, t+T-1$.

Lemma 2 *Suppose Assumptions 1 and 2 hold. For every* $\epsilon > 0$ *and* $T > (\log \frac{\epsilon}{2} - \log \bar{v}) / \log \delta$, *there is an* α *such that for every simple strategy* $\mathbf{y}$, *if* $\pi_t^{\mathbf{y}T} > 1 - \alpha$, *then in equilibrium* $\mu_t \in E_t^\epsilon(\mathbf{y}, \lambda)$ *for all* λ *and all* t.

Proof: In the case with no strategic externality to prove that $\mu_t \in E_t^\epsilon(\mathbf{y}, \lambda)$ it suffices to show that for all $(x, z) \in \operatorname{supp}\mu_t$, $x \in B_t^\epsilon(\mathbf{y}, z)$ (the aggregate action μ has been dropped as an argument of $B(.)$ since by Assumption 2 v is independent of μ).

Choose a T such that

$$\delta^T \bar{v} < \eta$$

or, taking logs,

$$T > \frac{\log \eta - \log \bar{v}}{\log \delta}.$$

Let $x_t \in B_t^\epsilon(\mathbf{y}, z)$ and $x'_t \notin B_t^\epsilon(\mathbf{y}, z)$. Let V_t be the expected payoff along the equilibrium path if x_t is chosen in period t and the player behaves optimally otherwise and let V'_t be the expected payoff along the equilibrium path if x'_t is chosen and the player behaves optimally otherwise. Then

$$V_t - V'_t \geq (1-\alpha)\epsilon - \alpha\bar{v} - \eta \tag{7}$$

Note that this inequality holds independent of the particular choice of x_t and x'_t. To show that an ϵ best response to $\mathbf{y}$ is played in equilibrium we need to show that there is an α such that $V_t - V'_t > 0$. From (7) a sufficient condition for that to happen is:

$$(1-\alpha)\epsilon - \alpha\bar{v} - \eta > 0 \tag{8}$$

For $\eta < \epsilon/2$ there is an α such that (8) is satisfied and the Lemma follows. □

Lemma 2 shows that if the small players believe that the large player will play a given sequence of actions for a sufficiently long period of time with a suficiently high probability, then they will play an aggregate ϵ best response to it. Lemma 1 on the other hand showed that if the large player played a certain strategy long enough then the small players would become convinced that he will continue to play that strategy for the following T periods with an arbitrarily high probability.

The following Lemma applies Lemma 1 and Lemma 2 to the Stackelberg strategy described above to show that in all but a finite number of periods the small players will play an aggregate ϵ best response to an optimal sequence if the large player imitates a Stackelberg type.

Let $T^* > (\log \frac{\epsilon}{2} - \log \bar{v})/\log \delta$ and let $\mathbf{y}^*$ denote the strategy played by commitment type $\omega(\epsilon, \eta, T^*)$. Let H^* be the set of histories consistent with $\mathbf{y}^*$ being played by b. Further let

$$H_t^*(\epsilon, \eta, \lambda) = \{h \in H_t^* | y = y(\epsilon, \eta, \lambda), \mu \in E^\epsilon(y(\epsilon, \eta, \lambda)), \lambda_1 = \lambda\}$$

be the histories for which the sequence $y(\epsilon, \eta, \lambda)$ and an aggregate ϵ best response to this sequence have been played.

Lemma 3 *Suppose $h \in H^*$. Then there is a number N, independent of h, such that the number of periods for which $\mu_t \notin E_k^\epsilon(y(\epsilon, \eta, \lambda), h_k)$, for all $h_k \in H_k^*(\epsilon, \eta, \lambda)$, for all $k \leq t$ and for all λ is bounded by N with probability 1.*

Proof: For the proof of this Lemma we keep (ϵ, η) fixed and therefore we will drop (ϵ, η) as arguments in $y(.)$ and $H^*(.)$. Suppose that for all $k \leq t$, $\mu_t \notin E_k^\epsilon(y(\lambda), h_k)$, $h_k \in H_k^*(\lambda)$. Then there is a $t' \in (t - T + 1, \ldots, t)$ such that $\pi_{t'}^{*T} < (1 - \alpha)$ since otherwise $h_t \backslash h_{t'} \in H_{t-t'}^*(\lambda_{t'})$ for some $0 \leq t' \leq t$ and hence the large player will continue to play $y(\lambda_{t'})$ for the next T periods with probability greater than $1 - \alpha$ and therefore Lemma 2 implies that $\mu_t \in E_k^\epsilon(y(\lambda), h_k)$.

But $\pi_t^{*T} < (1 - \alpha)$ at most $T\frac{\log p^*}{\log(1-\alpha^*)}$ times with probability 1 (Lemma 1). Thus $N \leq T^2 \frac{\log p^*}{\log(1-\alpha^*)}$ with probability 1. □

The following Lemma says that by imitating the commitment type constructed in the section 4 the large player can get a payoff at least $\bar{V}(\epsilon) - \eta$. Since ϵ and η are arbitrary Lemma 4 proves Theorem 1.

Lemma 4 *Suppose Assumptions 1 and 2 hold. Further suppose that $\omega^*(\epsilon, \eta)$ has prior probability $p^* > 0$ then in any Nash equilibrium $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ for initial state λ, $\lim_{\beta \to 1} V^b(\beta; \boldsymbol{\sigma}, \boldsymbol{\mu}, \lambda) \geq \bar{V}^b(\epsilon) - \eta$.*

Proof: Consider the strategy for b of always following y^* (corresponding to $\omega^* = \omega(\epsilon, \eta, T)$). Then $\mu_t \notin E_k^\epsilon(y(\epsilon, \eta, \lambda), h_k)$ for fewer than N periods by Lemma 3. For a given λ let

$$v^t(\lambda) = \inf_{\mu \in E^\epsilon(y(\lambda))} (1 - \beta) \sum_{k=1}^{t} \beta^{k-1} v(y_k(\epsilon, \eta, \lambda), \mu_k),$$

for $\mu_1^Z = \lambda$ and let $v^t = \inf_\lambda v^t(\lambda)$. Then a lower bound for b's Nash equilibrium payoff can be described as:

$$v^{t_1} + 0 + \beta^{t_1+1} v^{t_2} + \ldots + \beta^{t_1+\ldots+t_{N-1}+N-1} v^{t_N} + 0 + \beta^{t_1+\ldots+t_N+N} v^\infty \geq$$
$$\geq v^t + 0 + \beta^{t+1} v^t(\lambda) + + \ldots + \beta^{(N-1)t+N-1} v^t + 0 + \beta^{Nt+N} v^\infty$$

for some t where $0 \leq t \leq \infty$. Then

$$\begin{aligned} V^b(\beta; \boldsymbol{\sigma}, \boldsymbol{\mu}, \lambda) &\geq v^t(1 + \beta^{t+1} + \beta^{2(t+1)} + \ldots + \beta^{(N-1)(t+1)}) + \beta^{N(t+1)} v^\infty \\ &\geq v^{t+1}(1 + \beta^{t+1} + \beta^{2(t+1)} + \ldots + \beta^{(N-1)(t+1)}) + \beta^{N(t+1)} v^\infty \\ &\quad -\beta^{t+1}(1 - \beta)\bar{v}(1 + \beta^{t+1} + \beta^{2(t+1)} + \beta^{(N-1)(t+1)}) \\ &= v^{t+1} \frac{1 - \beta^{N(t+1)}}{1 - \beta^{t+1}} + \beta^{N(t+1)} v^\infty - \beta^{t+1}(1 - \beta)\bar{v} \frac{1 - \beta^{N(t+1)}}{1 - \beta^{t+1}} \end{aligned} \tag{9}$$

Let

$$\hat{v}^t = \frac{v^t}{(1 - \beta^t)} = \frac{\sum_{k=1}^t \beta^{k-1} v_k^t}{\sum_{k=1}^t \beta^{k-1}}$$

Then (9) becomes

$$V^b(\beta;\boldsymbol{\sigma},\boldsymbol{\mu},\lambda) \quad \geq \quad \hat{v}^{t+1}(1-\beta^{N(t+1)}) + \beta^{N(t+1)}v^{\infty} - \beta^{t+1}\bar{v}\frac{1-\beta^{N(t+1)}}{1-\beta^{t+1}}(1-\beta) \quad (10)$$

Now we want to let $\beta \to 1$. Notice that the t that appears in (10) is a function of β. If $t(\beta)$ stays bounded by some $T < \infty$ as $\beta \to 1$, then the first and the last term (10) tend to zero and the result follows since

$$\lim_{\beta\to 1} \beta^{N(t+1)}v^{\infty} \geq \bar{V}^b(\epsilon) - \eta.$$

If $t(\beta)$ does not stay bounded as $\beta \to 1$, i.e. $t(\beta) \to \infty$, then we have for some $0 \leq \theta \leq 1$:

$$\lim_{\beta\to 1} V^b(\beta;\boldsymbol{\sigma},\boldsymbol{\mu},\lambda) \geq (1-\theta)\liminf_{\beta\to 1} \hat{v}^{t(\beta)+1} + \theta \lim_{\beta\to 1} v^{\infty} \geq \bar{V}^b(\epsilon) - \eta$$

since both $\liminf_{\beta\to 1} \hat{v}^{t(\beta)+1}$ and $\lim_{\beta\to 1} v^{\infty}$ are greater than or equal to $V(\epsilon) - \eta$. □

9.3 Proof of Theorem 2

Theorem 2 *Suppose that Assumptions 1,3, and 5 hold. In any Nash equilibrium $(\boldsymbol{\sigma},\boldsymbol{\mu})$ for initial state λ, $\lim_{\beta\to 1} V^b(\beta;\boldsymbol{\sigma},\boldsymbol{\mu},\lambda) \geq \bar{V}$.*

First we will need a preliminary Lemma.

Lemma 0 *For all $\epsilon > 0$, there is a pair (η,ζ) such that if $\mu_1 \in E_1^{\zeta}(\boldsymbol{\sigma},\lambda)$ and $|\mu_1 - \mu_1'| < \eta$, then $\mu_1' \in E_1^{\epsilon}(\boldsymbol{\sigma},\lambda)$.*

Proof: $\mu_1 \in E_1^{\zeta}(\boldsymbol{\sigma},\lambda)$ means that there exists a $\boldsymbol{\mu} \in E^{\zeta}(\boldsymbol{\sigma}) : \boldsymbol{\mu}_1(h_1) = \mu_1$, where $h_1 = \lambda$. Moreover for each realization of $y \in Y^{\infty}$, $\boldsymbol{\mu}$ implies a $\mu \in M^{\infty}$. Clearly we can construct a $\boldsymbol{\mu}'$ such that for each realization of $y \in Y^{\infty}$ we get a $\mu' \in M^{\infty}$ with $|\mu' - \mu|_{\infty} < \eta$. By the definition of aggregate ϵ best response there exists a μ_1'' with $|\mu_1'' - \boldsymbol{\mu}_1(h_1)| < \zeta$ such that:

$$x \in B_1^{\zeta}(\boldsymbol{\mu},\boldsymbol{\sigma};h_1,z), \forall (x,z) \in \mathrm{supp}\mu_1''.$$

However, since $|\mu_1' - \mu_1| = |\mu_1' - \boldsymbol{\mu}_1(h_1)| < \eta$ we have $|\mu_1' - \mu_1''| < \eta + \zeta$. By continuity of v, for all $\epsilon > 0$ there exist ζ and η such that

$$x \in B_1^{\epsilon}(\boldsymbol{\mu}',\boldsymbol{\sigma};h_t,z), \forall (x,z) \in \mathrm{supp}\mu_1''$$

which means that $\mu_1' \in E_1^{\epsilon}(\boldsymbol{\sigma},\lambda)$. □

The next Lemma is a weaker version of Lemma 2 for the case where v depends on μ. Recall that for a pure strategy $\mathbf{y}$, $\pi_t^{\mathbf{y}T}$ denotes the probability that $\mathbf{y}$ is played in each of the next T periods, i.e. in the periods $t, t+1, t+2, \ldots, t+T$.

Lemma 5 *Given $\epsilon > 0$, for every L there is a (T, α) such that for $\pi_t^{yT} > 1 - \alpha$ then in equilibrium for all $y \in Y(L)$, $\mu_t \in E_t^\epsilon(y, h_t)$ where (T, α) is independent of t, h_t.*

Proof: Note that since $Y(L)$ contains a finite number of elements and since $(y_t, y_{t+1}, \ldots) \in Y(L)$ if $y \in Y(L)$ it is sufficient to show that for every pure strategy $\mathbf{y}$ we find a (T, α) such that if $\pi_1^{\mathbf{y}T} > 1 - \alpha$, then $\mu_1 \in E_1^\epsilon(\mathbf{y}, \lambda)$, for all λ.
Let $\Sigma^T(\mathbf{y}, \alpha) = \{\sigma | y_1, \ldots, y_T$ is played with probability $1 - \alpha\}$. If $\pi_1^{T,y} > 1 - \alpha$ then in equilibrium $\boldsymbol{\mu} \in E(\boldsymbol{\sigma}, \lambda)$ for some $\boldsymbol{\sigma} \in \Sigma^T(\mathbf{y}, \alpha)$.
Let $(\boldsymbol{\sigma}^T)$ be a sequence such that $\boldsymbol{\sigma}^T \in \Sigma^T(y, \alpha^T), \alpha^T \to 0$ as $T \to \infty$. Let $(\boldsymbol{\mu}^{*T}, \lambda^T)$ be a sequence such that $\boldsymbol{\mu}^{*T} \in E(\boldsymbol{\sigma}^T, \lambda^T)$ and let $\mu^* = \boldsymbol{\mu}(h^*)$ where h^* is the history where $\sigma_t = y_t$ for all t.

Claim: If μ^*, λ is a limit point of $(\mu^{*T}), \lambda^T$, then $\mu^* \in E(\mathbf{y}, \lambda)$.

Pf: Suppose $\mu^* \notin E(\mathbf{y}, \lambda)$. Then there is a τ and a set $D \subset Z \times X$ with $\sum_{(z,x) \in D} \mu^*(z, x) > \gamma > 0$ and for all $(z, x) \in D$

$$x \notin B_\tau(\mathbf{y}, \mu^*; h_\tau, z)$$

Thus if $(z, x) \in D$ then for all $\mathbf{x}$ with $\mathbf{x}(h_\tau) = x$ there exists an $\mathbf{x}'$ such that:

$$V_\tau(\mathbf{y}, \mu^*, \mathbf{x}, z, h_\tau) \leq V_\tau(\mathbf{y}, \mu^*, \mathbf{x}', z, h_\tau) - \eta.$$

Choose T such that

$$\delta^T \bar{v} \leq \eta/2$$

Since $\mu_{t'}^{*T} \to \mu_{t'}^*$ uniformly for $t' \leq \tau + T$ and since $\alpha^T \to 0$ as $\boldsymbol{\sigma}^T \to \mathbf{y}$, by continuity of v in μ and in α at $\alpha = 0$, it follows that for large T and for $(z, x) \in D$ there is an x' such that

$$V_\tau(\boldsymbol{\sigma}^T, \mu^{*T}, \mathbf{x}, z, h_\tau) \leq V_\tau(\boldsymbol{\sigma}^T, \mu^{*T}, \mathbf{x}', z, h_\tau) - \eta/4. \tag{11}$$

However, for large T, $|\mu_\tau^{*T} - \mu_\tau^*| < \gamma/2$ and hence (11) contradicts the fact that $\boldsymbol{\mu}^{*T} \in E(\boldsymbol{\sigma}^T, \lambda^T)$. •

Next we want to show that as $T \to \infty$ the distance between the first element of an aggregate best response to $\boldsymbol{\sigma}^T$, $\mu_1^T \in E_1(\boldsymbol{\sigma}^T, \lambda)$, and the set of first elements of the aggregate ζ best response to $\mathbf{y}$, $E_1^\zeta(\mathbf{y}, \lambda)$, tends to zero. More precisely, we want to show that for all (ζ, η) there exists a T^* such that for all $T > T^*$ for $M^T(\lambda) = \{\mu | \mu \in E_1(\boldsymbol{\sigma}^T, \lambda)$ for some $\boldsymbol{\sigma}^T \in$

$\Sigma(\mathbf{y}, \alpha^T)\}$

$$\sup_{\lambda} \sup_{\mu_1^T \in M^T(\lambda)} \inf_{\mu \in E_1^\zeta(\mathbf{y},\lambda)} |\mu_1^T - \mu| \leq \eta.$$

This is satisfied if

$$\limsup_{T\to\infty} \sup_{\lambda} \sup_{\mu_1^T \in M^T(\lambda)} \inf_{\mu \in E_1^\zeta(\mathbf{y},\lambda)} |\mu_1^T - \mu| = 0.$$

From above we know that μ^*, λ a limit point of a sequence μ^T, λ^T with $\mu^T \in E(\boldsymbol{\sigma}^T, \lambda^T)$ has to belong to $E(\mathbf{y}, \lambda) \subseteq E^\zeta(\mathbf{y}, \lambda)$ which implies that the limit above is zero.

This implies (by Lemma 0) that by choosing η and ζ appropriately, $\mu_1^T \in E_1^\epsilon(\mathbf{y}, \lambda)$ for $T > T^*$, for all λ. □

Proof of Theorem 2: Let $\omega(\eta, \epsilon) = \omega(\epsilon, \eta, T^*)$ denote the commitment type that plays the Stackelberg strategy described above, where T^* satisfies Lemma 5 uniformly for all $\mathbf{y} \in Y(L)$ and L is chosen sufficiently large so that there is a $y(\epsilon, \eta, \lambda) \in Y(L)$ for all λ. Now we can apply Lemma 3. Given that Lemma 3 holds so does Lemma 4. Note that η, ϵ can be chosen arbitrarily by Assumption 3. Thus Lemma 4 proves Theorem 2. □

9.4 Proof of Theorem 3

Theorem 3 *Suppose Assumptions 1, 2, 6, and 7 hold and all players have a common discount factor δ. Then in any Nash equilibrium $(\boldsymbol{\sigma}, \boldsymbol{\mu})$ for initial state λ, $\lim_{\delta\to 1} V^b(\delta; \boldsymbol{\sigma}, \boldsymbol{\mu}, \lambda) \geq \hat{V}^b$.*

Proof: *Step 1:* Let T be such that for all λ

$$\min_{\mu^T \in E^{T,\epsilon}(y^T(\epsilon,\lambda),\lambda)} \frac{1}{T}\sum_{t=1}^{T} v^b(y_t^T(\epsilon, \lambda), \mu_t) \geq \hat{V}^b(\lambda, \epsilon) - \eta.$$

Note that for all $\epsilon, \eta > 0$ there is a $T < \infty$ such that $y^T(\epsilon, \lambda)$ satisfies the above inequality *for all* λ. This is the case since

$$|\hat{V}^b(\epsilon, \lambda) - \hat{V}^b(\epsilon, \lambda')| < \bar{v} \cdot |\lambda - \lambda'|$$

since v is independent of μ[12].

Step 2: **Claim** *Let* $\mathbf{y}$ *be a given pure strategy. Independent of* $\mathbf{y}$, *for any* $\epsilon > 0$ *there are* $\alpha > 0$ *and* $\bar{\delta} < 1$ *and a* $T < \infty$, *such that if the probability that* $\mathbf{y}$ *is followed in the first* T *periods is greater than* $1 - \alpha$, *then for all* $1 \geq \delta \geq \bar{\delta}$ *in any Nash equilibrium* $(\mu_1, \dots, \mu_T) \in E^{T,\epsilon}(\mathbf{y}, \lambda_1)$.

[12] The constant $\bar{v}$ is the upper bound on the payoffs of the small and the large players (Assumption 1).

Pf: Let

$$B^{T,\epsilon}(\mathbf{y},z) = \{(x_t,z_t)_{t=1}^T \in G^T(y^T)|z_1 = z \text{ and for all } (x'_t,z'_t) \in G^T(y^T) \text{ with } z'_1 = z$$
$$\frac{1}{T}\sum_{t=1}^T v(y_t^T,x_t,z_t) \geq \frac{1}{T}\sum_{t=1}^T v(y_t^T,x'_t,z'_t) - \epsilon\}$$

Let

$$v^{*T}(z) = \frac{1}{T}\max_{\{x_t\}}\sum_{t=1}^T v(y_t,x_t,z_t)$$

with $z_1 = z$.

There are 3 reasons why a small player may not want to play an element in $B^{T,\epsilon}(\mathbf{y},z)$. First, $\mathbf{y}$ will only be followed with probability $1-\alpha$; second, playing a best response may cause the player to reach a state in period T which is not the optimal state for the play thereafter and third, the player discounts future payoffs, instead of using the time-average criterion.

Let $\sum_{t=1}^T \delta^{t-1}v_t$ be the expected payoff of the small player along the equilibrium path in the next T periods and let z_{T+1} be the state in which player i is in period $T+1$ along the equilibrium path. For $(x^T,z^T) \notin B^{T,\epsilon}(\mathbf{y},z)$ we have:

$$\sum_{t=1}^T \delta^{t-1}v_t \leq (1-\alpha)(v^{*T}(z)-\epsilon) + \alpha\bar{v} + \bar{v}\cdot\sum_{t=1}^T\left|\frac{1}{T} - \delta^{t-1}\frac{1-\delta}{1-\delta^T}\right|$$

On the other hand, the player can use the following sequence: for the first $T-N$ periods play a sequence that maximizes the average payoff in the first T periods against y, in the last N periods, adjust the state so that in period $T+1$ the state z_{T+1} is reached. This gives a lower bound on the payoff:

$$\sum_{t=1}^T \delta^{t-1}v_t \geq (1-\alpha)v^{*T} - \bar{v}\cdot\sum_{t=1}^T\left|\frac{1}{T} - \delta^{t-1}\frac{1-\delta}{1-\delta^T}\right| - \bar{v}\frac{N}{T}$$

Now note that if T is large and α is close to zero and δ is close to one then the prescribed strategy is an element in $B^{\epsilon,T}(\mathbf{y},z)$. Furthermore it gives a larger payoff than any strategy that is not an element of $B^{\epsilon,T}(y^T)$ since

$$(1-\alpha)\epsilon > 2\bar{v}\cdot\sum_{t=1}^T\left|\frac{1}{T} - \delta^{t-1}\frac{1-\delta}{1-\delta^T}\right| + \bar{v}\frac{N}{T} + \alpha\bar{v}$$

But this implies that for all $(z_t,x_t)_{t=1}^T \in \text{supp}(\mu_1,\ldots,\mu_T)$ such that $z_{t+1} = f(y_t,z_t,x_t)$ we have $(z_t,x_t)_{t=1}^T \in B^{\epsilon,T}(\mathbf{y},z_1)$, which proves the claim. •

Step 3: Let μ denote the sequence of (μ_t) induced by the history when player b imitates the type $\omega(\epsilon, T)$ and let π_t^{*T} be the probability that $\mathbf{y}(\epsilon, T)$ is played in the periods $t, t+1, \ldots, t+T-1$. For every $\alpha > 0$, $\pi_{kT+1}^{*T} < 1 - \alpha$ for fewer than $N(\alpha, T)$ different k (Lemma 1). Thus for all but $N(\alpha, T)$ different k we have $(\mu_{kT+1}, \ldots, \mu_{kT+T}) \in E^{\epsilon, T}(y^T(\epsilon, \lambda_{kT+1}), \lambda_{kT+1})$. But this implies that for all but N periods of length T the undiscounted payoff of b is larger than $\hat{V}^b(\epsilon) - \eta$. Since ϵ, η can be chosen arbitrarily small (Assumption 7), the Theorem follows. $\square$

References

[1] Ausubel L. M. and R. Deneckere, (1989), "Reputation in Bargaining and Durable Goods Monopoly", *Econometrica* 57, 511-531.

[2] Becker, G. (1975), *Human Capital: A Theoretical and Empirical Analysis*, Columbia University Press, New York, NY, for NBER, 2nd edition.

[3] Chari, V.V. and P. Kehoe (1990), "Sustainable Plans", *Journal of Political Economy*, **98**, 783-802.

[4] Coase, R. (1972), "Durability and Monopoly", *Journal of Law and Economics*, **15**, 143-149.

[5] Cripps, M., and J. Thomas (1992), "Reputation and Commitment in Two-Person Repeated Games", mimeo, University of Warwick.

[6] Dekel, E. and J. Farell (1990), "One-Sided Patience with One-Sided Communication Does Not Justify Stackelberg Equilibrium", *Games and Economic Behavior*, **2**, 299-303.

[7] Dutta, P. (1991), "A Folk Theorem for Stochastic Games", mimeo, Columbia University.

[8] Fischer, S. (1980), "Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government ", *Journal of Economic Dynamics and Control*, **2**, 93-107.

[9] Fudenberg, D., and D. Levine (1989), "Reputation and Equilibrium Selection in Games with a Single Patient Player", *Econometrica* **57**, 759-78.

[10] Fudenberg, D., and D. Levine (1992), "Maintaining a Reputation when Strategies are Unobserved", *Review of Economic Studies*, **59**(3), 561-580.

[11] Fudenberg, D., D. Levine, and J. Tirole (1985), "Infinite Horizon Models of Bargaining with One-Sided Incomplete Information", in A. Roth (ed.) *Game Theoretic Models of Bargaining*, Cambridge University Press.

[12] Gul, F., H. Sonnenschein, and R. Wilson, (1986), "Foundations of Dynamic Monopoly and the Coase Conjecture", *Journal of Economic Theory*, **39**, 155-190.

[13] Kreps,D., and R. Wilson, (1982) "Reputation and Imperfect Information", *Journal of Economic Theory* **27**, 253-79.

[14] Kydland F., and E. Prescott (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans", *Journal of Political Economy*, **85**, 473-91.

[15] Kydland F., and E. Prescott (1980), "Dynamic Optimal Taxation, Rational Expectations and Optimal Control", *Journal of Economic Dynamics and Control*, **2**, 79-91.

[16] Milgrom, P., and J. Roberts (1982), "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory* **27**, 280-312.

[17] Ramsey, F. P. (1927), "A Contribution to the Theory of Taxation", *Economic Journal* **37**, 47-61.

[18] Schmidt, K. (1992), "Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interest", MIT working paper 92-7.

[19] Stokey N. (1981) "Rational Expectations and Durable Goods Pricing", *Bell Journal of Economics,* 12, 112-128.

[20] Stokey, N. (1992) "Credible Public Policy", *Journal of Economic Dynamics and Control*, **15**, 627-656.