

Reputation in Dynamic Games*



MARCO CELENTANI

Universidad Carlos III de Madrid, Madrid, 28903, Spain

AND

WOLFGANG PESENDORFER

Northwestern University, Evanston, Illinois 60208

Received May 3, 1993; revised July 23, 1995

We analyze reputation in a game between a large player and a continuum of long-lived small players in which state variables affect players' payoffs. The large player's type is private information. We give conditions under which in every Nash equilibrium a very patient large player will get almost the largest payoff consistent with the small players choosing a best response in a large finite truncation of the game. While our results apply to the time inconsistency problem of optimal government policy, we show that for the durable goods monopoly reputation may fail to improve the monopolist's payoff. *Journal of Economic Literature* Classification Numbers: C72, C73.

1. INTRODUCTION

We consider dynamic games with one larger player and a large number (continuum) of small players. Large and small players are long lived and their payoffs are affected by state variables whose evolution is described by a possibly stochastic transition law. As is standard in reputation models, we suppose that the large player's type is private information and that he may be a "commitment type" who is locked into playing a particular strategy.

* Marco Celentani gratefully acknowledges the financial support of DGICYT, Grant PB92-0245 and of HCM Programme, grants ERBCHICT 940975 and CHRXCT 940458. Wolfgang Pesendorfer gratefully acknowledges the financial support of the National Science Foundation, Grant SBR-9409180. We wish to thank an associate editor, two anonymous referees, V. V. Chari, Eddie Dekel, David Levine, Andy Newman, Asher Wolinsky, and seminar participants at UCLA, Northwestern University, Universidad Carlos III de Madrid, and Universitat Pompeu Fabra for useful discussions and suggestions. All remaining errors are ours.

The purpose of the paper is to investigate the extent to which an uncommitted or “normal” type of the large player can exploit his opponents’ uncertainty to establish a reputation for a particular behavior.

One application of this model is a policy game in which a government faces a large number of households (see Fischer [10] and Chari and Kehoe [3]). The government must decide, for example, on capital and labor taxes and households choose investment and labor supply. A second application is the analysis of a market for a durable good in which there is a monopoly seller and a continuum of potential buyers (see Coase [4], Gul *et al.* [15]).

Previous work on reputation has considered repeated games and in most studies it was assumed that a long-lived player faces a short-lived (or myopic) opponent (see Kreps and Wilson [17], Milgrom and Roberts [22], Fudenberg and Levine [12, 13]). We are interested in a situation in which all players are long-lived but we will assume that the game is *anonymous* in the sense that a deviation by a single small player does not affect the public history of the game.¹

In a repeated game our anonymity assumption implies that each small player will play a short-run best response in each period to that period’s expected play since his actions do not affect his future payoffs or the public history of the game. Consequently, in a repeated game the best possible commitment for the long-run player is to the Stackelberg strategy for the corresponding static game. Moreover, the long-run player only needs to convince his effectively myopic opponents that he will follow the Stackelberg strategy in the current period in order to achieve the Stackelberg payoff for that period.

The case of dynamic games differs from that of repeated games in two ways. First, the optimal commitment strategy will generally not be a constant action in every period. Second, the small players’ current play may depend on expected future play of the large player. As a consequence the large player may have to demonstrate that he follows the commitment strategy not only in the current period but also for a large number of future periods. For example, suppose that the large player is a government deciding whether to place a tax on capital or on labor and small players are households who can decide how much to invest. If capital does not depreciate in one period, households will choose a high investment level today only if they are convinced that the government will set low capital tax rates for a sufficient number of future periods.

Our analysis will be carried out in two parts. In Section 3 we derive a lower bound on equilibrium payoffs of an arbitrarily patient large player

¹ As Levine and Pesendorfer [20] demonstrate, anonymous equilibria can be obtained as the limit of equilibrium outcomes in games with a large, finite number of small players and noisy observability of each small player’s action.

playing against small players having a fixed discount factor strictly smaller than 1. Theorem 1 shows that in every Nash equilibrium an arbitrarily patient large player's payoff is bounded below by the average payoff he could get by optimal commitment in a large finite truncation of the game when the initial state is the least favorable among all the states that can be reached under this commitment strategy.

We apply Theorem 1 to a stylized version of a capital taxation model and to a game illustrating the durable goods monopoly problem. In both cases, the payoff bound is tight, i.e., the large player receives the largest payoff consistent with small players behaving optimally.

Theorem 1 does not provide a tight payoff bound in games in which small players can get "trapped" in states that give the large player a low payoff. Suppose that the large player wishes to commit to a strategy which leads to a very undesirable state if the small players do not play a best response to that strategy. In this case reputation arguments do not work since by the time small players learn that the large player is following this commitment strategy it may be too late for them to avoid the undesirable state. We give an example of a trade liberalization game, similar to Matsuyama [21], in which this issue arises.

The second part of the analysis is presented in Section 4, where we consider the case in which the large and the small players are equally patient. This case is particularly relevant for policy games, in which the payoff function of the government is typically related to the payoff functions of private agents. Theorem 2 shows that the large player will be able to exploit his reputation if the transition function satisfies the following *reversibility* condition. A transition function exhibits reversibility if for every strategy of a small player and for every deviation from that strategy there is a return strategy that leads to the path of the original strategy or to a better path within a bounded number of periods.

The reversibility condition is satisfied in capital accumulation games, but is not satisfied in the standard durable goods monopoly. Once a customer has purchased the durable good, he has reached an irreversible state. We give an example that demonstrates how in the durable goods monopoly reputational arguments fail to guarantee the large player his optimal commitment payoff.

Our analysis is related to the literature on reputation with two long-run players. Schmidt [24] and Cripps *et al.* [7] analyze two-player repeated games with one patient and one non-myopic but less patient player and show how incorrect off-equilibrium-path beliefs can weaken reputation effects. By introducing imperfect observability of actions, Celentani *et al.* [5] ensure that the less patient player can learn the full strategy of the patient player and in turn that the latter can establish a reputation for playing strategies that use rewards and punishments.

The difference between the current study and the studies on reputation with two long-lived players is that our anonymity assumption implies that an individual small player cannot be deterred from playing a best response to their beliefs about equilibrium-path play of other players. This enables us to extend our results to the case in which both players are arbitrarily patient.

Our work is also related to the literature on dynamic games. Dutta [9] provides a Folk theorem for stochastic games. Chari and Kehoe [3] and Stokey [25] study the time inconsistency problem introduced by Kydland and Prescott [18, 19] and Fischer [10] and characterize the set of equilibria in problems of optimal policy design when the government cannot commit.

Both Chari and Kehoe [3] and Stokey [25] show that, if there is sufficiently little discounting, a desirable outcome (the Ramsey outcome in a capital taxation model, Ramsey [23]) can arise in equilibrium. However, in their model the Ramsey outcome is only one of many equilibria² while our results suggest that it may be the unique Nash equilibrium outcome if reputational arguments are used.

In addition, we give an example of a capital taxation model (Section 2.1) for which the optimal commitment outcome cannot be achieved as a Nash equilibrium of the complete information game even if all players are very patient. Nevertheless, if we introduce commitment types in this example, our results show that a patient government will achieve the optimal commitment payoff.

2. DESCRIPTION OF THE GAME

We consider a game with one large player and a continuum of identical small players $i \in [0, 1] = I$. The finite sets Y and X denote the sets of actions of the large player and the small players, respectively; $y \in Y$, $x \in X$. Let R denote the set of mixed actions for the large player and let S denote the set of mixed actions for a small player. Each small player is characterized by his state, $z \in Z$, where Z is finite. Let \mathcal{A} denote the set of probability measure on Z , $\lambda \in \mathcal{A}$, and let 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 in state z that choose action x .

² It is sometimes argued that in this case the government may be able to select its preferred equilibrium. Dekel and Farrell [8], however, show that these selection arguments are inconsistent.

The game is played in the following way: At the beginning of each period $t = 1, 2, \dots$ 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 chooses a mixed action $s(i) \in S$ and the large player simultaneously chooses a mixed action $r \in R$. After these actions have been selected, payoffs occur and all players observe the realization of the mixed action of the large player y_t and the distribution μ_t of actions of the small players; μ_t is a joint distribution over actions and states, i.e., after each period every player learns the measure of small players in state z whose mixed action had a realization x , for every $(z, x) \in Z \times X$. Note however, that μ_t does not reveal the play of any individual small player since μ_t is unaffected by deviations of sets of players of measure zero.

The law of motion for the state variables of the small players is described by the function $f: Y \times M \times X \times Z \rightarrow A$. Thus $\lambda_{t+1}(i) = f(y_t, \mu_t, x_t(i), z_t(i))$ describes the probability distribution over the states of player i in period $t + 1$, given aggregate play and player i 's action and state in period t . We assume that f is continuous on M . Note that if $z_t(i)$ and $s_t(i)$ are measurable functions then $\lambda_{t+1}(i)$ is a measurable function. By $\lambda_t = \int_I \lambda_t(i) di$ we denote the aggregate state³ in period t , where $\lambda_t(z)$ denotes the fraction of small players that are in state z in period t .

The initial state of the game is a measurable function $z_0: I \rightarrow Z$. We assume that z_0 is common knowledge for all players. The public history of the game at time t consists of the sequence of realizations of $y_\tau, \mu_\tau, \tau = 1, \dots, t - 1$. For notational convenience the history at time t also includes $\lambda_\tau, \tau = 1, \dots, t$, where $\lambda_1 = \int_I z_0(i) di$.⁴ The set of histories in period t is denoted by $H_t = (Y \times M \times A)^{t-1} \times A$, 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 $t > 1$. The set of infinite histories will be denoted by $H = H_\infty$. For the history from t' to $t, t' < t$, we write $h_t \setminus h_{t'} \in H_{t-t'}$. A private history for a small player is a sequence of states and actions. The set of private histories in period t is denoted by $\tilde{h}_t = X^{t-1} \times Z^{t-1} \times Z$.

For a given sequence of play $(y, \mu) = ((y_t, \mu_t)_{t=1}^\infty)$ the payoff to the large player is

$$(1 - \beta) \sum_{t=1}^{\infty} \beta^{t-1} w(y_t, \mu_t).$$

³ Implicitly in the definition of the aggregate state we assume that the strong law of large numbers holds with a continuum of random variables (see Judd [16]). However, as Levine and Pesendorfer [20] show for a different context, the model with a continuum of small players can be interpreted as the limit of a model with a finite number of players and noisy observation of the aggregate play of the small players.

⁴ Notice that given the transition law, λ_t is determined by μ_{t-1}, y_{t-1} .

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

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} v(y_t, \mu_t, x_t, z_t).$$

This notation allows us to describe types of small players with different payoff functions by an appropriate choice of state variables and the initial state. We illustrate this in the example describing the durable goods monopoly below.

Assumption 1. $w, v,$ and f are continuous on M . Moreover $0 \leq v, w \leq \bar{v}$.

A mixed (behavioral) strategy for the large player in period t is a mapping $\rho_t: H_t \rightarrow R$. Similarly a mixed strategy for the small players in period t is a measurable function $\sigma_t: [0, 1] \times H_t \times \tilde{h}_t \rightarrow S$, $\sigma = (\sigma_t)_{t=1}^\infty$. We assume throughout this paper that the strong law of large numbers holds for a continuum of i.i.d. random variables. Thus, given a strategy σ of the small players and a sequence of action (y_t) for the large player a unique sequence of public outcomes (y_t, μ_t) is induced.⁵

Let $\bar{\sigma}_t: H_t \rightarrow M$ denote an *aggregate strategy* for the small players in period t , $\bar{\sigma} = (\bar{\sigma}_t)_{t=1}^\infty$. We say that $\bar{\sigma}$ is an aggregate strategy for σ and the initial state z_0 if for any sequence $(y_t) \in Y^\infty$, $((y_t), \sigma, z_0)$ and $((y_t), \bar{\sigma}, \int_I z_0(i) di)$ induce the same public history h . Note that if $\bar{\sigma}$ is an aggregate strategy for (σ, z_0) then $\bar{\sigma}(h_t)$ is uniquely determined for all public histories h_t for which all but a measure zero subset of small players have followed $\sigma_{t'}(i)$ for all $t' < t$. For histories that are reached following a deviation of a positive measure of small players the aggregate strategy can be arbitrary. This is done for notational convenience since histories that follow simultaneous deviations of small players are irrelevant.

The payoff of every player depends on his own strategy, on the large player's action, and on the aggregate strategy of the small players. We denote by $V(\delta, \rho, \bar{\sigma}, \sigma(i), h_t, \tilde{h}_t(i))$ the expected payoff to player i from playing $\sigma(i)$, after history $(h_t, \tilde{h}_t(i))$ if the aggregate strategy of the small players is $\bar{\sigma}$. Similarly $W(\beta, \rho, \bar{\sigma}, h_t)$ is the expected payoff to the large player after history h_t if the aggregate strategy of the small players is $\bar{\sigma}$.

⁵ The aggregate outcome in period 1 is defined as $\mu_1(x, z) = \int_{\{i: z_0(i)=z\}} \sigma_1(i)(x) di$. Suppose that μ_1, \dots, μ_{t-1} are determined. Then for every small player i , $z_0(i)$, $(\sigma_1(i), \dots, \sigma_t(i))$ together with μ_1, \dots, μ_{t-1} and y_1, \dots, y_{t-1} induce a probability distribution g_t^i on M , where $g_t^i(x, z)$ is the probability that agent i is in state z and chooses action x in period t . Moreover, note that g_t^i is a measurable function of i since $z_0(i)$ and $\sigma(i)$ are measurable. Thus we can define $\mu_t = \int_I g_t^i di$.

2.1. Example 1

This example is a stylized version of the classic time inconsistency problem of government policy (see Fischer [10], Chari and Kehoe [3]). Suppose the large player is a government that must choose whether to place a tax on capital income or use a distortionary labor tax in an effort to raise adequate revenue. Small players are households who decide how much of their current endowment to consume and how much to invest. The government decides on taxes: y^z and y^l are the capital and labor tax rates respectively. Consumers are born with no capital ($z_1 = 0$) and have a labor income of 1 each period (before taxes). If a consumer invests $1/2$ units of his income, then he has one unit of capital in the following period ($z = 1$). Furthermore, capital does not depreciate and any further investment does not increase the capital stock. One unit of capital produces one unit of income every period. Thus the consumer's actions are given by $X = \{0, 1/2\}$. The consumer's utility in period t is given by:⁶

$$v(y_t, x_t, z_t) = z_t(1 - y_t^z) + (1 - y_t^l)c - x_t,$$

where $c - 1 > 0$ is the cost of imposing a distortionary labor tax.⁷

Government utility is equal to the sum of the households' utility, minus a penalty in the case the revenue is different from 1. Let λ_t^1 denote the fraction of households who are in state $z = 1$. Let μ_t^x denote the fraction of households who choose $x = 1/2$ in period t . Then,

$$w(y_t, \mu_t) = \lambda_t^1(1 - y_t^z) + (1 - y_t^l)c - \mu_t^x/2 - P \cdot |\lambda_t^z \lambda_t^1 + y_t^l - 1|.$$

We assume that $P > c > 1$ and that the government is at least as patient as the households, i.e., $\beta \geq \delta$. It is straightforward to show that $\lambda_t^1 = 0$ is the *unique* Nash equilibrium outcome⁸ for all $\beta \geq \delta \in [0, 1)$. Even though for $\delta > 1/3$ a social planner would want to invest and reach state $\lambda^1 = 1$, this state cannot be achieved as an equilibrium.

⁶ For simplicity, instead of a budget constraint, we assume that the consumer has a direct disutility of investing.

⁷ As is standard in capital taxation models (see, e.g., Chari and Kehoe [3]) the government can enforce capital taxes even though it cannot observe individual states. One justification for this assumption is that there is an enforcement agency that can actually observe individual states whereas the agent who makes the policy decision has no access to this information. Alternatively, we can assume that the large player can observe individual histories but his strategy only depends on the aggregate history of the game. In this way we select a subset of the set of Nash equilibria. Fudenberg *et al.* [14] show that this equilibrium selection can be justified if with small probability the small players' chosen action is different from the observed action.

⁸ For a proof see Celentani and Pesendorfer [6].

2.2. Example 2

The following example is a simple version of the durable goods monopoly. 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. The large player is the seller of the good. Every period the monopolist sets a price y_t where $y_t \in \{0, 1/n, \dots, (5n-1)/n\}$; $n \geq 6$. If the buyer did not previously purchase the good he is in state $z = (0, H)$ if he is a high type and state $z = (0, L)$ if he is a low type. If the buyer is in state 0 in period t he can either take action 0 (he does not buy) or action 1 (he buys). If the buyer buys in period t at a price y then he moves to "state" (y, j) , $j = H, L$, where he stays forever (i.e., each sale is "final") enjoying a per period payoff of $(1 - \delta)(r^j - y)$.

Thus the transition function is defined as

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

We assume that the marginal cost of producing the good is zero. Therefore, if μ_t^1 denotes the fraction of buyers who purchase in period t , the payoff to the large player in period t is $w_t(y_t, \mu_t) = y_t \cdot \mu_t^1$.

Suppose that $\delta = \beta$. In this case Gul *et al.* [15] show that for large n and for δ close to one the payoff for the monopolist in any subgame perfect equilibrium is close to 2.

2.3. Example 3

The following example is a modified version of the trade liberalization game analyzed by Matsuyama [21].

A government (large player) faces a large number of firms (small players) which can be in one of three states, $Z = \{0, 1, 2\}$. A firm in state 1 is competitive on the international market whereas a firm in state 0 is not. A firm in state 2 is out of business. All firms are initially in state 0. Firms in state 0 can invest ($x = 1$) or not invest ($x = 0$). If a firm in state 0 invests in period t then from period $t + 1$ on it is in state 1. The government can liberalize in the current period ($y = 0$) or not liberalize ($y = 1$). If the government liberalizes in period t then all firms who are in state 0 are driven out of business and hence move to state 2 and stay there forever after. The government's per period payoff is zero if it does not liberalize and is equal to the fraction of firms who are in state 1 if it liberalizes. The idea is that the government prefers to liberalize but it also wants its firms to survive international competition. Firms' payoffs are

$$v(y, x, z) = \begin{cases} -x & \text{if } z = 2 \text{ or } y = 1, z = 0 \\ 1 - x & \text{if } z = 1 \\ 1 - x & \text{if } z = 0, y = 0. \end{cases}$$

Clearly, it is a subgame perfect equilibrium of this game for the firms never to invest and for the government never to liberalize.

3. REPUTATION AND EQUILIBRIUM PAYOFFS

In the first part of this section we derive a benchmark for the payoff of the large player when he can commit to a strategy in large finite truncations of the game.

Fix a strategy ρ . Suppose the large player plays the following “restarting” strategy: he follows ρ for T periods; in period $T + 1$ the large player restarts strategy ρ , i.e., he follows ρ for the initial state λ_{T+1} from period T on; in period $2T + 1$ he again restarts ρ , and so on. The set $A(\rho, \lambda)$ is the set of aggregate states that can be reached if the initial aggregate state is λ and the large player uses a restarting strategy based on ρ with no additional restrictions imposed on small players’ strategies.⁹

DEFINITION 1. $A(\rho, \lambda) \equiv \bigcup_{k=1}^{\infty} A_k(\rho, \lambda)$, where $A_k(\rho, \lambda)$ is defined as follows:

(i) $\lambda' \in A_1(\rho, \lambda)$ if there exists a (σ, z_0) with $\int_I z_0(i) di = \lambda$ such that if (ρ, σ) is played and the initial state is z_0 then there is a t such that the aggregate state in period t is λ' .

(ii) For $k > 1$, $\lambda' \in A_k(\rho, \lambda)$ if there exists a (σ, z_0) with $\int_I z_0(i) di \in A_{k-1}(\rho, \lambda)$ such that if (ρ, σ) is played and the initial state is z_0 then there is a t such that the aggregate state in period t is λ' .

In Example 1 above, $A(\rho, \lambda) = A$ for all (ρ, λ) , since every state can be reached independent of the actions of the large player. In Example 2, however, the actions of the large player determine the set of aggregate states that can be reached. If, for example, the large player chooses a strategy according to which the price is never below 5, then $A(\rho, \lambda)$ contains only the initial aggregate state together with all aggregate states in which some of the buyers move from their initial state to state $y = 5$.

In the following, when we define the payoff a large player can guarantee himself by committing to ρ we assume that the worst aggregate state in $A(\rho, \lambda)$ is realized.

⁹ Recall that a strategy pair (σ, ρ) together with an initial state z_0 induce a deterministic sequence of aggregate states (λ_t) .

The T -period expected average payoff of the large player if $(\rho, \bar{\sigma})$ are played and the initial state is λ is defined as

$$W^T(\rho, \bar{\sigma}, \lambda) = \frac{1}{T} E \left[\sum_{t=1}^T w(y_t, \mu_t) \mid \rho, \bar{\sigma}, \lambda \right].$$

Similarly, for small player i we define the discounted payoff in period t of a T -period truncation of the game as

$$\begin{aligned} & V_i^T(\delta, \rho, \bar{\sigma}, \sigma(i), h_t, \tilde{h}_t(i)) \\ &= \frac{1-\delta}{1-\delta^{T-t+1}} E \left[\sum_{t'=t}^T \delta^{t'-t} w(y_{t'}, \mu_{t'}, x_{t'}, z_{t'}(i)) \mid \rho, \bar{\sigma}, \sigma(i), h_t, \tilde{h}_t(i) \right]. \end{aligned}$$

Define $V_i^T(1, \cdot) = \lim_{\delta \rightarrow 1} V_i^T(\delta, \cdot)$ to be the average (undiscounted) payoff in period t of a T -period truncation of the game.

By $B^T(\varepsilon, \delta, \rho, \lambda)$ we denote the aggregate strategies of the small players that are generated if the small players choose an ε best response to ρ in every period of a T -period truncation of the game with the exception of the last ε fraction of periods.

DEFINITION 2. $\bar{\sigma} \in B^T(\varepsilon, \delta, \rho, \lambda)$ if there exists a strategy σ and an initial state z_0 such that:

- (1) $\lambda = \int_I z_0(i) di$ and $\bar{\sigma}$ is an aggregate strategy for (σ, z_0) .
- (2) For all $t = 1, \dots, T - \tau$ with $\tau < T\varepsilon$ and for all $i \in I$, $V_i^T(\delta, \rho, \bar{\sigma}, \sigma(i), h_t, \tilde{h}_t(i)) \geq V_i^T(\delta, \rho, \bar{\sigma}, \sigma'(i), h_t, \tilde{h}_t(i)) - \varepsilon$ for all $\sigma'(i)$ and for all $h_t, \tilde{h}_t(i)$ that are reached with strictly positive probability under $(\rho, \bar{\sigma}, \sigma(i))$ when the initial state is $(\lambda, z_0(i))$.

In the following we define $W^*(\delta, \lambda)$ as the set containing all undiscounted (average) payoffs that the large player can get in large finite truncations of the game if he commits to a pure strategy ρ and if the worst aggregate state $\lambda' \in A(\rho, \lambda)$ and the worst strategy in $B^T(\varepsilon, \delta, \rho, \lambda')$ are realized.

DEFINITION 3. $w \in W^*(\delta, \lambda)$ if for every $\eta > 0$ there is an $\varepsilon > 0$ such that for every $T' < \infty$ there is a $T > T'$ and a pure strategy ρ with $W^T(\rho, \bar{\sigma}, \lambda') \geq w - \eta$ for all $\lambda' \in A(\rho, \lambda)$ and for all $\bar{\sigma} \in B^T(\varepsilon, \delta, \rho, \lambda')$.

The goal of the rest of this section is to show that if there is incomplete information about the type of the large player then the Nash equilibrium payoff of a patient large player is bounded below by the largest element (supremum) in the set $W^*(\delta, \lambda)$ if λ is the initial aggregate state of the game (Theorem 1).

We assume that the large player can be one of countably many types $\omega \in \Omega$ and that these types are drawn from a common knowledge prior p assigning a strictly positive probability to all points in Ω . The large player is informed of his type before play begins, but this is purely private knowledge and is not revealed to the small players. The large player's payoff function now also depends on his type, $W(\beta, \rho, \bar{\sigma}, \omega, h_t)$. We will focus on a particular type $\omega_0 \in \Omega$ which we refer to as the rational type. Payoffs for the rational type are as defined in the previous section, i.e., $W(\beta, \rho, \bar{\sigma}, \omega_0, h_t) \equiv W(\beta, \rho, \bar{\sigma}, h_t)$.

A *type behavior strategy* for the large player specifies a time-indexed sequence of maps from public histories to mixed actions for that type. We denote these by $\rho(\omega)$. Since the small players cannot observe the type of the large players, the definition of a strategy for the small players remains unchanged.

A *Nash equilibrium* for initial state z_0 is a $((\rho(\omega))_{\omega \in \Omega}, \sigma)$ such that if $\bar{\sigma}$ is an aggregate strategy for (σ, z_0) , $\int_I z_0(i) di = \lambda$, then for all $\omega \in \Omega$, $W(\beta, \rho(\omega), \bar{\sigma}, \omega, \lambda) \geq W(\beta, \rho'(\omega), \bar{\sigma}, \omega, \lambda)$ for all $\rho'(\omega)$ and all $\omega \in \Omega$ and

$$\begin{aligned} & \sum_{\omega \in \Omega} p(\omega) V(\delta, (\rho(\omega)), \bar{\sigma}, \sigma(i), \lambda, z_0(i)) \\ & \geq \sum_{\omega \in \Omega} p(\omega) V(\beta, (\rho(\omega)), \bar{\sigma}, \sigma'(i), \lambda, z_0(i)) \end{aligned}$$

for all $i \in I$ and for all $\sigma'(i)$.

Remark. An alternative interpretation of the game is that the individual histories of small players are publicly observable but that equilibrium strategies only depend on public histories (and on own individual histories for small players). Under this interpretation we select a subset of Nash equilibria in which small players are strategically anonymous.¹⁰

It is easy to show by taking limits of finite truncations of this infinite game that Nash equilibria exist. See, for example, Fudenberg and Levine [11].

Next we define a collection of commitment types which are assumed to have strictly positive prior probability. These types are the analog of the "Stackelberg type" in the literature on reputation in repeated games (see Fudenberg and Levine [12]). Given δ and the initial state λ we denote by $\rho^*(T, n, \delta, \lambda)$ a pure strategy which gives the large player an average payoff that is within $1/n$ of the supremum of all payoffs that he can achieve by

¹⁰ Fudenberg, Levine, and Pesendorfer [14] show that if each small player's action is imperfectly observable (even if this imperfection is very small) then the equilibrium of a game with a finite number of small players converges to an equilibrium in which small players are strategically anonymous as the number of small players goes to infinity.

committing to a pure strategy in a T -period truncation of the game subject to small players playing a $1/n$ best response with the possible exception of the last $1/n$ fraction of periods. More precisely, $\rho^*(T, n, \delta, \lambda)$ satisfies

$$\begin{aligned} & \inf_{\lambda' \in \mathcal{A}(\rho^*, \lambda)} \inf_{\bar{\sigma} \in B^T(1/n, \delta, \rho^*, \lambda')} W^T(\rho^*, \bar{\sigma}, \lambda') \\ & \geq \sup_{\rho \in \mathcal{Y}} \inf_{\lambda' \in \mathcal{A}(\rho, \lambda)} \inf_{\bar{\sigma} \in B^T(1/n, \delta, \rho, \lambda')} W^T(\rho, \bar{\sigma}, \lambda') - 1/n, \end{aligned} \quad (1)$$

where \mathcal{Y} denotes the set of pure strategies for the large player.

Now we define the strategy $\rho(T, n, \delta)$ as follows: In period $t = kT + \tau$, $1 \leq \tau \leq T$, $k = 0, 1, \dots$,

$$\rho_t(T, n, \delta)(h_t) = \rho_\tau^*(T, n, \delta, h_1)(h_t \setminus h_{kT}), \quad (2)$$

where $h_1 \equiv \lambda_1$ is the aggregate state in the beginning of the game. Let $\omega(T, n, \delta)$ denote the type that plays the strategy $\rho(T, n, \delta)$, or in other words “restarts” the strategy $\rho^*(T, n, \delta, h_1)$ every T periods. Note that the set of commitment types $\{\omega(T, n, \delta), T = 1, 2, \dots; n = 1, 2, \dots\}$ is a countable set and therefore we can assume that each type has a strictly positive prior probability.

Let $N_t(\beta, \delta, \lambda)$ denote the least (inf) expected payoff to the large player conditional on type ω_0 in any Nash equilibrium if the initial aggregate state is λ . We are now in the position to state our first result.

THEOREM 1. *Suppose that Assumption 1 is satisfied and that $\omega(T, n, \delta) \in \Omega$ for all (T, n) . Then $\liminf_{\beta \rightarrow 1} N_t(\beta, \delta, \lambda) \geq \sup W^*(\delta, \lambda)$.*

Theorem 1 says that in every Nash equilibrium a patient large player gets at least the time average of payoffs corresponding to an optimal commitment in a large finite truncation of the game.

The intuition for this result can be summarized as follows. First, suppose that the small players believe that the large player will play a particular strategy for T future periods with very high probability. We show that in this case behavior of the small players in the first T periods implies that they choose an ε best response for a T period truncation of the game provided that T is sufficiently large. In a second step we show that if the large player imitates any commitment type that has positive prior probability then there is at most a finite number of periods in which the small players believe that the probability that the large player will play the commitment strategy in the following T periods is bounded away from 1. Finally, we can split the game into T -period “superstages,” i.e., the first

T periods are the first superstage, the second T periods are the second superstage, and so on. Since ε and T were arbitrary, the large player can guarantee himself a payoff arbitrarily close to the best payoff in $W^*(\delta, \lambda)$ in all but a finite number of “abnormal” superstages. In the limit, as $\beta \rightarrow 1$, the abnormal superstages do not affect the large player’s payoff and the theorem follows.

3.1. The Examples with Reputation

3.1.1. *Example 1.* In the time-consistency example, when $\lambda_1 = 0$, the unique Nash equilibrium payoff of the government is $1 - c < 0$. For the perturbed game we can apply Theorem 1 to compute a lower bound for all Nash equilibrium payoffs. As can be verified easily, if $\delta > 1/3$, $\sup W^*(\delta, \lambda) = 1$ and hence a patient large player gets a payoff of almost 1 in every Nash equilibrium of the perturbed game.

The following strategy is an example for a commitment strategy that achieves this payoff. As long as the aggregate capital stock is below 1 the government rewards investment with zero taxes on capital for a number of periods which is just sufficient to give households an incentive to invest. Thus the commitment strategy is $y_t^z = 0$ for all t such that $\hat{\lambda}_{t-\tau} < 1$ and $y_t^z = 1$ otherwise, where τ is the smallest integer that satisfies $\delta(1 - \delta^\tau)/(1 - \delta) > 1/2$. For large finite truncations of the game the average payoff of the large player is 1 if he plays this commitment strategy and the small players choose an ε best response in all but the last ε fraction of periods.

3.1.2. *Example 2.* In the durable goods monopoly case, the monopolist can guarantee himself a payoff close to the total consumer surplus in the perturbed game. To see this note that the following strategy is a possible commitment strategy: Choose $p_t = (5n - 1)/n$ for τ periods. If half of the consumers purchase by period τ then the monopolist sets a price of $(2n - 1)/n$ for $t' > \tau$ on. Otherwise the monopolist chooses a price of $(5n - 1)/n$ for all $t' = \tau + 1, \dots, 2\tau$, and so on. If τ is chosen so that $\delta^\tau(5 - (2n - 1)/n) < 1/n$, an ε best response to this strategy in any truncated game with sufficiently large T and sufficiently small ε implies that all consumers whose evaluation is 5 purchase in the first period and all other consumers purchase in period $\tau + 1$. Moreover, the only aggregate states that can be reached under this strategy are states in which either all purchases are made at a price of $(5n - 1)/n$ or all high types purchase at a price of $(5n - 1)/n$. Therefore, $\sup W^*(\delta, \lambda) = 7/2 - 2/(2n)$, which is close to $7/2$ for large n .

3.1.3. *Example 3.* The simple trade liberalization example highlights the difference between being able to commit to a strategy and establishing a reputation for playing a particular strategy in a dynamic game. For Example 3, Theorem 1 does not imply a payoff larger than the worst

subgame perfect equilibrium in the unperturbed game. To see this note that for any policy for which “investing” is an ε best response for the firms (for small ε) the government must choose a policy ρ according to which “liberalization” is chosen in some period. But this implies that $A(\rho, \lambda) = A$, and therefore the aggregate state in which every firm has gone out of business is an element in $A(\rho, \lambda)$. Clearly, if all firms have gone out of business the payoff to the government is zero and hence $\sup W^*(\delta, \lambda) = 0$. The government cannot exploit a reputation in this example since in order to build up a reputation it has to liberalize. Liberalization, on the other hand, will lead to an undesirable state if the firms do not play a best response to the commitment strategy. Since there is no reason why firms should *always* play a best response to the commitment strategy, the government cannot benefit from establishing a reputation.

3.2. Proof of Theorem 1

The proof of Theorem 1 uses two preliminary lemmas. In Lemma 1 we investigate the consequences of imitating a particular commitment type on the beliefs of the small players and 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 extensions of a result in Fudenberg and Levine [12].

Let ρ be the pure strategy played by a particular commitment type ω^* . Let h^* be the event that $y_t = \rho_t(h_t)$ for all h_t that are reached following (ρ, σ) starting from a given $h_1 = \lambda_1$. Furthermore let $p(\omega^*) = p^*$ denote the prior probability of type ω^* . Let $\pi_t^\tau(\rho)$ be the probability that in the next τ periods the actions of the large player are consistent with ρ , i.e., the probability that in periods $t, t+1, \dots, t+\tau-1$ aggregate play is consistent with ρ being played, $\pi_t^\tau(\rho) = \Pr[y_t = \rho_t(h_t), \dots, y_{t+\tau-1} = \rho_{t+\tau-1}(h_{t+\tau-1}) | h_{t-1}, \sigma]$. Finally, let $n(\pi_t^\tau(\rho) \leq \bar{\pi})$ be the random variable denoting the number of periods in which $\pi_t^\tau(\rho) \leq \bar{\pi}$.

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

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

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.

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

$$\begin{aligned} & \Pr(\omega^* | h_{t+1}) \\ &= \Pr(\omega^* | \rho_t(h_t), h_t) \\ &= \frac{\Pr(\omega^* | h_t) \Pr(\rho_t(h_t) | \omega^*)}{\Pr(\omega^* | h_t) \Pr(\rho_t(h_t) | \omega^*) + (1 - \Pr(\omega^* | h_t)) \Pr(\rho_t(h_t) | \bar{\omega}^*)}. \end{aligned} \quad (3)$$

Notice that $\Pr(\rho_t(h_t) | \omega^*) = 1$ and that the denominator of (3) is equal to $\Pr(\rho_t(h_t))$. Therefore (3) can be rewritten as

$$\Pr(\omega^* | h_{t+1}) = \frac{\Pr(\omega^* | h_t)}{\Pr(\rho_t(h_t))}. \quad (4)$$

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

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

Suppose that $\pi_t^\tau(\rho) \leq \bar{\pi}$ for all $t' = t, \dots, t + \tau - 1$. Then if the large player plays like the commitment type for $t' = t, \dots, t + \tau - 1$ (i.e., $y_{t'} = \rho_{t'}(h_{t'})$), the probability that he is type ω^* has to go up by a factor of at least $1/\bar{\pi}$ (because if $y_{t'} = \rho_{t'}(h_{t'})$ for all $t' = t, \dots, t + \tau - 1$, then at some $t' = t, \dots, t + \tau - 1$, $\bar{\pi}_t^\tau(\rho)$ 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 $\pi_t^\tau(\rho) \leq \bar{\pi}$ for τK periods during which $y_t = \rho(h_t)$, in every period then $\Pr(\omega^* | h_{K \cdot \tau+1}) \geq p^*/\bar{\pi}^K$. However, since

$$\Pr(\omega^* | h_t) \leq 1 \quad (5)$$

if

$$p^*/\bar{\pi}^K > 1, \quad (6)$$

inequality (5) is violated and a contradiction to the hypothesis that $\pi_t^\tau(\rho) \leq \bar{\pi}$ for all $t' = t, \dots, t + K \cdot \tau - 1$ is obtained. Taking the log of (6) the condition becomes $K > \log p^*/\log \bar{\pi}$ and the proof is complete. ■

The following lemma shows that if the small players believe that with high enough probability the large player will play a given strategy in a large enough truncation of the game, then they will play an ε best response to it.

LEMMA 2. *Suppose Assumption 1 is satisfied. For any $\varepsilon > 0$, $\delta < 1$ there are $\alpha > 0$ and $\bar{T} < \infty$, such that, for a fixed pure strategy ρ and an initial aggregate state λ , if the probability that ρ is followed in the first $T \geq \bar{T}$ periods is greater than $1 - \alpha$, then if the small players play a best response, $\bar{\sigma} \in B^T(\varepsilon, \delta, \rho, \lambda)$. Moreover α and \bar{T} are independent of ρ and λ .*

Proof. Let $V_i^{*T}(i)$ be the maximal discounted payoff of a small player in period t if the game ends in period T and $(\rho, \bar{\sigma})$ is being played. (The maximum is taken over the individuals' action x_1, \dots, x_T .) If $\bar{\sigma} \notin B^T(\varepsilon, \delta, \rho, \lambda)$ then there is a period $1 \leq t \leq T - \tau$ and an i such that the expected payoff in period t is bounded above by

$$(1 - \delta^{T-t+1})((1 - \alpha)(V_i^{*T}(i) - \varepsilon) + \alpha \bar{v}) + \delta^{T-t+1} \bar{v}.$$

On the other hand, by playing a $\sigma(i)$ that is a best response in the T -period truncation player i gets at least $(1 - \delta^{T-t+1})(1 - \alpha) V_i^{*T}(i)$. Note that

$$\begin{aligned} & [(1 - \delta^{T-t+1})((1 - \alpha)(V_i^{*T}(i) - \varepsilon) + \alpha \bar{v}) + \delta^{T-t+1} \bar{v}] \\ & \quad - [(1 - \delta^{T-t+1})(1 - \alpha) V_i^{*T}(i)] \\ & < (1 - \delta^\tau)(1 - \alpha)(-\varepsilon) + \alpha \bar{v} + \delta^\tau \bar{v} < 0 \end{aligned}$$

for τ sufficiently large and α sufficiently close to zero. In addition we can choose T so that $\tau/T < \varepsilon$ and hence the lemma follows. ■

Proof of Theorem 1. Choose an n^* such that for all \bar{T} there are a $T > \bar{T}$ and a $\rho^*(T, n^*, \delta, \lambda)$ that satisfy

$$\inf_{\lambda' \in A(\rho^*, \lambda)} \inf_{\bar{\sigma} \in B^T(1/n^*, \delta, \rho^*, \lambda')} W^T(\rho^*, \bar{\sigma}, \lambda') \geq \sup W^*(\delta, \lambda) - \eta. \quad (7)$$

Note that for all $\eta > 0$ and for all λ there is an n^* such that (7) is satisfied (this follows from the definition of $W^*(\delta, \lambda)$). Now we can choose $T^* > \bar{T}$ sufficiently large and $\alpha^* > 0$ sufficiently small so that Lemma 1 is satisfied for $(\varepsilon^* = 1/n^*, \alpha^*, T^*)$.

Suppose the large player imitates type $\delta(T^*, n^*, \delta)$, who is committed to $\rho(T^*, n^*, \delta)$ and whose prior probability is $p^* > 0$. Even though we do not claim that imitating type $\omega(T^*, n, \delta)$ is part of an equilibrium, the large player's equilibrium payoff cannot be smaller than the payoff the large player could get by playing such a strategy. In the following we compute

a lower bound on the payoff the large player could get by playing this strategy, which will in turn provide a lower bound on his equilibrium payoffs.

Let $\pi_t^{T^*}$ be the probability that $\rho(T^*, n^*, \delta)$ is played in the periods $t, t+1, \dots, t+T^*-1$. Let $N^* = T^* \log p^* / \log(1 - \alpha^*)$. Lemma 1 implies that $\pi_t^{T^*} < 1 - \alpha^*$ for fewer than N^* periods, which in turn implies that $\pi_{kT^*+1}^{T^*} < 1 - \alpha^*$ for fewer than N^* different k . Thus for all but at most N^* different k we have $\bar{\sigma}(h_{kT^*+1}) \in B^{T^*}(1/n^*, \delta, \rho^*, \lambda_{kT^*+1})$, which in turn implies that for all but at most N^* different k

$$W^{T^*}(\rho(T^*, n^*, \delta), \bar{\sigma}, \lambda_{kT^*+1}) \geq \sup W^*(\delta, \lambda) - \eta.$$

Therefore if the large player imitates type $\omega(T^*, n^*, \delta)$ his payoff is at least

$$\beta^{N^*T^*} (\sup W^*(\delta, \lambda) - \eta).$$

Since η can be chosen arbitrarily small the theorem follows. ■

4. PATIENT SMALL PLAYERS

In Theorem 1 we assume that the discount factor of the small players is bounded away from one while the large player is 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"). Moreover, a standard motivation for player's patience is a shortening of the interval between dates. In this case, if all players are long-lived, all discount factors have to tend to 1 at the same rate. For these reasons 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 stems from the fact that small players may become increasingly reluctant to take an action that leads to a particular 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 small players to enter a particular irreversible state.

For the sake of simplicity in this section we specialize to a deterministic transition function; formally we assume that if $\lambda(i) = f(y, \mu, x, z)$ then $\lambda(i)(z) \in \{0, 1\}$.¹¹

¹¹ A similar result can be obtained for the case of a stochastic transition function.

DEFINITION 4. State \hat{z} dominates state z given $(\rho(\omega)_{\omega \in \Omega}, \bar{\sigma}, \lambda)$ if there is a $\hat{\sigma}(i)$ and a $\bar{\delta} < 1$ such that

$$V(\bar{\delta}, \rho(\omega), \bar{\sigma}, \sigma(i), \lambda, z) \leq V(\bar{\delta}, \rho(\omega), \bar{\sigma}, \hat{\sigma}(i), \lambda, \hat{z})$$

for all $\delta \in (\bar{\delta}, 1)$, for all $\sigma(i)$, and for all $\omega \in \Omega$.

Let $z(h_t, \sigma(i), z)$ be the state reached in period t starting from z if the public history is h_t and i plays the pure strategy $\sigma(i)$. Let $Z(h_t, z) = \{z \mid z = z(h_t, \sigma(i), z) \text{ for some } \sigma(i)\}$.

DEFINITION 5. z' and z'' are related if $\{z', z''\} \subset Z(h_t, z)$ for some (h_t, z) .

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

Assumption 2 (Reversibility). There is a number $N < \infty$ with the following property: Suppose that z' and z'' are related. Fix any $(\rho(\omega)_{\omega \in \Omega}, \bar{\sigma})$ and any $\sigma(i)$. Then there is a pure strategy $\tilde{\sigma}(i)$ such that $z(h_t, \tilde{\sigma}(i), z'')$ dominates $z(h_t, \sigma(i), z')$ given $(\rho(\omega)_{\omega \in \Omega}, \bar{\sigma}, \lambda)$ for some $t \leq N$ with probability 1.¹²

Assumption 2 implies that whenever a small player deviates from the equilibrium path then for every state he reaches along this deviation there is a “return strategy” that allows the player to move back to the path induced by the equilibrium strategy (or to a dominating path) irrespective of the aggregate play.

Notice that the reversibility assumption is trivially satisfied in repeated games. Reversibility is also satisfied if the transition function only depends on the actions of the large player because in this case z' and z'' are related if and only if $z' = z''$.

Suppose that we can define a complete ordering on the states Z , such that if $z < z'$ then z' dominates z . Further suppose that if \hat{z} is the maximal state then \hat{z} can be reached from all states for all actions of the large player. In this case, Assumption 2 is satisfied. Note that this structure corresponds to the case where the state variable represents human or physical capital and the large player (government) cannot prevent the small players from accumulating capital. Thus Assumption 2 is satisfied in Example 1.

A second class of examples in which Assumption 2 is satisfied is the case where the transition can be expressed as a “chain.” Suppose that we can order the states in Z as z^0, z^1, \dots, z^K , where z^{k-1}, z^k, z^{k+1} are the only states that can be reached from z^k for $2 \leq k \leq K-1$, z^0 and z^1 are the only

¹² The probability distribution over histories is induced by $((\rho(\omega))_{\omega \in \Omega}, \bar{\sigma}, \lambda)$. Notice that N is independent of $z', z'', (\rho(\omega))_{\omega \in \Omega}, \bar{\sigma}$, and $\sigma(i)$.

states that can be reached from z^0 , and z^K and z^{K-1} are the only states that can be reached from z^K and moreover that the large player's action cannot prevent any of the feasible transitions in the sense that for every mixed action r there exists an x that guarantees that the small player is in state z' in the next period if z' is in the set of states that can be reached from today's state. In this case Assumption 2 is satisfied with $N = K$.

Assumption 2 is violated whenever the small player can reach a state z which does not dominate all other states and from which the small player cannot move away. In the durable goods monopoly example every purchase corresponds to a movement to a permanent state. Moreover, since purchasing at a low price is better than purchasing at a high price no strictly positive price dominates all other states.

Our reversibility assumption is related to the assumption of "asymptotic state independence" in Dutta [9] which requires the set of long-run feasible and individually rational payoffs to be independent of the initial state for a folk theorem to hold for stochastic games. In Dutta [9] this is needed in order to ensure that individual deviations do not unilaterally affect future payoff opportunities. In a similar fashion the reversibility assumption in our context guarantees that a deviation of an individual small player does not permanently reduce future payoff opportunities for that player.

THEOREM 2. *Suppose Assumptions 1 and 2 are satisfied and that $\omega(T, n, 1) \in \Omega$ for all (T, n) . Then $\liminf_{\delta \rightarrow 1} N_I(\delta, \delta, \lambda) \geq \sup W^*(1, \lambda)$.*

The idea behind the proof of Theorem 2 is that we split up the infinite game into finite superstages of length T . Note that the effect of a current decision of a small player on the payoffs in future superstages 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 active 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.

If the reversibility assumption is violated it is possible to construct counterexamples to Theorem 2. Below we give a counterexample for the durable goods monopoly problem based on Example 2.

4.1. *The Examples Revisited*

Example 1 satisfies Assumption 2 and hence the conclusion from the previous section is unchanged. The equilibrium payoff of a patient government is close to the best payoff the government can achieve. More generally, Assumption 2 is satisfied in the framework considered by Chari and Kehoe [3]. Hence we can conclude for this class of models that if both

the government and the households are very patient then in *any* Nash equilibrium the problem of time inconsistency of optimal government policy disappears in the perturbed game. In this case reputation effects allow the government to achieve a payoff that is equal to the payoff it could achieve by committing to an optimal policy.

Assumption 2 is violated in Example 2 (Section 2.2). If both the monopolist and the buyers share a common discount factor δ , then there exists a sequential equilibrium in which the monopolist gets a payoff of $2(n-1)/n$ for all δ . Note that if the monopolist could commit to an optimal policy then his payoff would be equal to the simple monopoly payoff of $5/2$.

To see that there is an equilibrium in the perturbed game which gives the monopolist a payoff of $2(n-1)/n$, suppose that 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 $\varepsilon > 0$. Note that type ω^* satisfies the definition of $\omega(T, n, 1)$ for all n and all T .

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 be shown that the normal type does not have an incentive to imitate type ω^* . Suppose the large player deviates and offers $p_t = (5n-1)/n$. Since

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

type H will not buy until period $T+1$.¹³ 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 the monopolist from deviating and the last is the

¹³ 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 $t=1$.

payoff from setting $p_1 = (2n - 1)/n$. This implies that a 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 of $(5n - 1)/2n$.

Remark. In the durable goods monopoly example we assume that the monopolist cannot repurchase the durable good from the consumer. Note, however, that Assumption 2 is violated even if we allow the monopolist to repurchase the good. This is the case since the small players's ability to return to state 0 depends on the *price* the monopolist is willing to pay when he repurchases the good. As long as we do not allow the monopolist to *commit* to a particular repurchase price when he sells the good the above example is robust to the introduction of repurchase by the monopolist. Only if the monopolist could commit to a repurchase price would the reversibility condition be satisfied. In this case the model would be strategically equivalent to a model in which the monopolist can rent the durable goods to the buyers and, in accordance with existing literature (see Bulow [2]), the monopolist can get the simple monopoly payoff.

4.2. Proof of Theorem 2

The following Lemma shows that if the small players believe that with high enough probability the large player will play a given strategy in a large enough truncation of the game, then they will play an ε best response to it provided that the reversibility assumption is satisfied.

LEMMA 3. *Suppose Assumptions 1 and 2 are satisfied. Let ρ be a given pure strategy and let λ be a given aggregate state. For any $\varepsilon > 0$ there are $\alpha > 0$, $\bar{\delta} < 1$ and $\bar{T} < \infty$, such that if the probability that ρ is followed in the first $T \geq \bar{T}$ periods is greater than $1 - \alpha$, then for all $1 \geq \delta \geq \bar{\delta}$ if the small players play a best response, $\bar{\sigma} \in B^T(\varepsilon, 1, \rho, \lambda)$. Moreover, α , $\bar{\delta}$ and \bar{T} are independent of ρ and λ .*

Proof. There are three reasons why small players may not want to play an element in $B^T(\varepsilon, 1, \rho, \lambda_1)$. First, ρ will be followed with a probability which is not sufficiently high; 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; third, the player discounts future payoffs, instead of using the time-average criterion.

Fix $(\rho, \bar{\sigma})$, let $V_t^T(\delta, h_t)(i)$ be the expected payoff of player i along this path in the $T - t$ periods starting after h_t and let $V_t^{*T}(1, h_t)(i)$ be the maximal average payoff for a small player along this path if the game ends in period T (the maximum is taken over the individual's action x_1, \dots, x_T). Let $z_{T+1}(i)(h_{T+1})$ be the state which is reached in period $T + 1$ along this path if h_{T+1} is the history and if $\sigma(i)$ is a pure strategy or the state which

is reached with strictly positive probability under h_{T+1} which maximizes the continuation payoff if $\sigma(i)$ is a mixed strategy.

If $\sigma \notin B^T(\varepsilon, 1, \rho, \lambda_1)$ then for some $1 \leq t' \leq T - \tau$, $\tau < T\varepsilon$ and for some i ,

$$V_{i'}^T(1, h_t)(i) \leq (1 - \alpha)(V_{i'}^{*T}(i) - \varepsilon) + \alpha\bar{v}.$$

On the other hand, every small player can use the following strategy: for the first $T - N$ periods play a strategy that maximizes the average payoff in the first T periods against $(\rho, \bar{\sigma})$; in the last N periods, adjust the state so that in period $t \leq T + 1$ the state $\hat{z}_t(h_t)$ is reached, where $\hat{z}_t(h_t)$ dominates $z_t(h_t)$. Assumption 2 guarantees that such a return strategy starting in period $T - N + 1$ exists. (Simply consider the subform starting in period $T - N + 1$. For every state reached in period $T - N$, Assumption 2 guarantees that there exists a $\hat{\sigma}$ such that for some $t \leq T + 1$ a state $\hat{z}_t(h_t)$ is reached that dominates $z_t(h_t)$.)

This gives a lower bound on $V_{i'}^T(1, h_t)(i)$, $1 \leq t' \leq T - \tau$:

$$V_{i'}^T(1, h_t)(i) \geq (1 - \alpha) V_{i'}^{*T}(i) - \bar{v} \frac{N}{\tau}.$$

If

$$(1 - \alpha) \varepsilon > \bar{v} \frac{N}{\tau} + \alpha\bar{v}$$

small player i 's average payoff increases when he switches to the prescribed strategy. Recalling that $\tau < T\varepsilon$, it is easy to see that this inequality can be satisfied if T is sufficiently large and α is sufficiently close to zero. Since $V_{i'}^T(\delta, h_t)(i)$ converges to $V_{i'}^T(1, h_t)(i)$ uniformly for all h_t we can choose a $\bar{\delta}$ sufficiently close to one so that the conclusion holds for all $\delta > \bar{\delta}$ and hence the lemma follows. ■

Proof of Theorem 2. The rest of the proof of Theorem 2 is omitted as it is identical to the proof of Theorem 1 once we replace δ with 1, Lemma 2 with Lemma 3, and β with δ .

5. CONCLUSIONS

In this paper we investigate reputation in dynamic games with a large player and a large number of small players whose individual actions are not observed. We consider two different environments:

(1) The small players are less patient than the large player, i.e., the small players' discount factor is bounded away from 1 while the large player is arbitrarily patient.

(2) Both the large and the small players are arbitrarily patient and the transition function satisfies a reversibility condition which ensures that a deviation of an individual small player does not permanently reduce future payoff opportunities for that player.

For both cases we show that every Nash equilibrium the large player gets a payoff close to the largest payoff he could get by optimally committing to a pure strategy in a finite truncation of the game if the initial aggregate state is the least favorable among the states that can be reached under the commitment strategy.

When the small players' discount factor is fixed, we show that reputation is a substitute for commitment in the classical time-inconsistency and durable goods monopoly problems. An example of a trade liberalization game shows how reputation may fail to have equilibrium implications if a state can be reached that prevents the large player from gaining from reputation.

In the case in which the large and the small players are arbitrarily patient, we show that reputation is a substitute for commitment if the transition function is such that no strategy of a small player has irreversible consequences on his payoff opportunities. This condition fails to be satisfied in the durable goods monopoly case (in which purchase leads to an absorbing state) and an example shows that in this case reputational arguments fail.

Throughout the paper we assume that the transition between aggregate states is deterministic and that commitment types only use pure strategies. Both these assumptions can be relaxed. In particular, our results can be extended to allow for reputation for mixed strategies and stochastic transition of the aggregate state by adopting a framework similar to Fudenberg and Levine [12]. Since in this case the link between the large players' choices and the observed outcome is stochastic the large player is subject to moral hazard. Even imperfect observations, however, provide some information about the large player's choices, and therefore an analog for the "learning lemma" (Lemma 1) can be found also in this case along the lines of Theorem 4.1 in Fudenberg and Levine [13]. In general the possibility of commitment to a mixed strategy will lead to an improved payoff bound in a similar way as in the case of repeated games.

REFERENCES

1. L. AUSUBEL AND R. DENECKERE, Reputation in bargaining and durable goods monopoly, *Econometrica* **57** (1989), 511–531.
2. J. BULOW, Durable goods monopolist, *J. Polit. Econ.* **90** (1982), 314–332.
3. V. V. CHARI AND P. KEHOE, Sustainable plans, *J. Polit. Econ.* **98** (1990), 783–802.

4. R. COASE, Durability and monopoly, *J. Law Econ.* **15** (1972), 143–149.
5. M. CELENTANI, D. FUDENBERG, D. LEVINE, AND W. PESENDORFER, Maintaining a reputation against a patient opponent, *Econometrica*, forthcoming.
6. M. CELENTANI AND W. PESENDORFER, Reputation in dynamic games, CMSEMS Discussion Paper No. 1009, Northwestern University, 1992.
7. M. CRIPPS, K. SCHMIDT, AND J. THOMAS, Reputation in perturbed repeated games, *J. Econ. Theory*, forthcoming.
8. E. DEKEL AND J. FARRELL, One-sided patience with one-sided communication does not justify Stackelberg equilibrium, *Games Econ. Behav.* **2** (1990), 299–303.
9. P. DUTTA, A folk theorem for stochastic games, *J. Econ. Theory* **66** (1995), 1–32.
10. S. FISCHER, Dynamic inconsistency, cooperation and the benevolent dissembling government, *J. Econ. Dynam. Control* **2** (1980), 93–107.
11. D. FUDENBERG AND D. K. LEVINE, Subgame perfect equilibria of finite and infinite horizon games, *J. Econ. Theory* **31** (1983), 251–268.
12. D. FUDENBERG AND D. K. LEVINE, Reputation and equilibrium selection in games with a single patient player, *Econometrica* **57** (1989), 759–778.
13. D. FUDENBERG AND D. LEVINE, Maintaining a reputation when strategies are unobserved, *Rev. Econ. Stud.* **59** (1992), 561–580.
14. D. FUDENBERG, D. LEVINE, AND W. PESENDORFER, When are non-anonymous players negligible?, CMSEMS Discussion Paper No. 1114, 1995.
15. F. GUL, H. SONNENSCHN, AND R. WILSON, Foundations of dynamic monopoly and the Coase conjecture, *J. Econ. Theory* **39** (1986), 155–190.
16. K. JUDD, The law of large numbers with a continuum of IID random variables, *J. Econ. Theory* **35** (1985), 19–25.
17. D. KREPS AND R. WILSON, Reputation and imperfect information, *J. Econ. Theory* **27** (1982), 253–279.
18. F. KYDLAND AND E. PRESCOTT, Rules rather than discretion: The inconsistency of optimal plans, *J. Polit. Econ.* **85** (1977), 473–491.
19. F. KYDLAND AND E. PRESCOTT, Dynamic optimal taxation, rational expectations and optimal control, *J. Econ. Dynam. Control* **2** (1980), 79–91.
20. D. K. LEVINE AND W. PESENDORFER, When are agents negligible?, *Amer. Econ. Rev.* **85** (1995), 1162–1170.
21. K. MATSUYAMA, Perfect equilibria in a trade liberalization game, *Amer. Econ. Rev.* **80** (1990), 480–492.
22. P. MILGROM AND J. ROBERTS, Predation, reputation, and entry deterrence, *J. Econ. Theory* **27** (1982), 280–312.
23. F. P. RAMSEY, A contribution to the theory of taxation, *Econ. J.* **37** (1927), 47–61.
24. K. SCHMIDT, Reputation and equilibrium characterization in repeated games of conflicting interests, *Econometrica* **61** (1993), 325–352.
25. N. STOKEY, Credible public policy, *J. Econ. Dynam. Control* **15** (1992), 627–656.