



HAL
open science

Simple bounds on the value of a reputation

Olivier Gossner

► **To cite this version:**

Olivier Gossner. Simple bounds on the value of a reputation. *Econometrica*, 2011, 79 (5), pp.1627-1641. 10.3982/ECTA9385 . halshs-00654683

HAL Id: halshs-00654683

<https://shs.hal.science/halshs-00654683>

Submitted on 29 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic
Theory in its Relation to Statistics and Mathematics*

<http://www.econometricsociety.org/>

Econometrica, Vol. 79, No. 5 (September, 2011), 1627–1641

SIMPLE BOUNDS ON THE VALUE OF A REPUTATION

OLIVIER GOSSNER

*Paris School of Economics, 75014 Paris, France and London School of
Economics, London, U.K.*

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website <http://www.econometricsociety.org> or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

SIMPLE BOUNDS ON THE VALUE OF A REPUTATION

BY OLIVIER GOSSNER¹

We introduce entropy techniques to study the classical reputation model in which a long-run player faces a series of short-run players. The long-run player's actions are possibly imperfectly observed. We derive explicit lower and upper bounds on the equilibrium payoffs to the long-run player.

KEYWORDS: Reputation, repeated games, incomplete information, relative entropy.

1. INTRODUCTION

THE BENEFITS OF BUILDING A REPUTATION can be measured by considering the equilibrium payoffs of repeated games with incomplete information on a player's preferences. In the benchmark model of [Fudenberg and Levine \(1989, 1992\)](#), there is uncertainty on the type of a long-run player, who is facing a succession of short-run players. The long-run player can either be of a simple commitment type, whose preferences are to play a fixed strategy at every stage of the game, or of a normal type, whose preferences are fixed and who is acting strategically.

We provide explicit lower and upper bounds on equilibrium payoffs to the long-run player of a normal type. For patient players, upper bounds converge to an optimistic measure of the long-run player's Stackelberg payoff in the one-shot game, while lower bounds converge to a pessimistic measure of this Stackelberg payoff if commitment types have full support.

The main technical contribution of this paper is to introduce a tool from information theory called *relative entropy* to measure the difference between the short-run player's predictions on his next signal conditional on the long-run player's type and unconditional on it. The major advantage of relative entropies is the chain rule property, which allows derivation of an explicit bound on the expected sum of these differences that depends only on the distribution of types. This bound is then used to derive bounds on the long-run player's equilibrium payoffs.

2. MODEL AND MAIN RESULT

2.1. *The Repeated Game*

We recall the classical reputation model ([Fudenberg and Levine \(1992\)](#), [Mailath and Samuelson \(2006\)](#)) in which a long-run player 1 interacts with a succession of short-lived players 2. In the stage interaction, the finite set of

¹The author is grateful to Drew Fudenberg for encouraging this project, and to Mehmet Ekmekci, Larry Samuelson, and Saturo Takahashi for insightful comments. Useful suggestions from an editor and two anonymous referees are gratefully acknowledged.

actions to player i is A_i . Given any finite set X , $\Delta(X)$ represents the set of probability distributions over X , so that the set of mixed strategies for player i is $S_i = \Delta(A_i)$.

The set of types to player 1 is $\Omega = \{\tilde{\omega}\} \cup \hat{\Omega}$. While $\tilde{\omega}$ is player 1's *normal* type, $\hat{\Omega} \subseteq S_1$ is a countable (possibly finite) set of *commitment types*. Each type $\hat{\omega} \in \hat{\Omega}$ is a simple type committed to playing the corresponding strategy $\hat{\omega} \in S_1$ at each stage of the interaction. The initial distribution over player 1's types is $\mu \in \Delta(\Omega)$ with full support.

Actions in the stage game are imperfectly observed. Player i 's set of signals is a finite set Z_i . When actions $a = (a_1, a_2)$ are chosen in the stage game, the pair of private signals $z = (z_1, z_2) \in Z := Z_1 \times Z_2$ is drawn according to the distribution $q(\cdot|a) \in \Delta(Z)$. For $\alpha \in S_1 \times S_2$, we let $q(z|\alpha) = \mathbf{E}_\alpha q(z|a)$, and $q^2(\cdot|\alpha)$ denotes the marginal distribution of $q(\cdot|\alpha)$ on Z_2 . Each player i is privately informed of the component z_i of z . Particular cases of the model include public monitoring ($z_1 = z_2$ almost surely for every a) and perfect monitoring ($z_1 = z_2 = a$ almost surely for every a). Each instance of player 2 is informed of the sequence of signals received by previous players 2. We do not impose that player 2's signal reveals player 2's action, thus allowing for the interpretation that a signal to player 2 is a trace, such as feedback on a website, that player 2 leaves to all future instances of other players 2.

The stage payoff function to player 1's normal type is u_1 , and player 2's payoff function is u_2 , where $u_i: A_1 \times A_2 \rightarrow \mathbb{R}$. The model includes the cases in which player i 's payoffs depend only on the chosen action and received private signal or on the private signal only.

The set of private histories prior to stage t for player i is $H_{i,t-1} = Z_i^{t-1}$, with the usual convention that $H_{i,0} = \{\emptyset\}$. A strategy for player 1 is a map²

$$\sigma_1: \Omega \times \bigcup_{t \geq 0} H_{1,t} \rightarrow S_1$$

that satisfies $\sigma_1(\hat{\omega}, h_{1,t}) = \hat{\omega}$ for every $\hat{\omega} \in \hat{\Omega}$, since commitment types are required to play the corresponding strategy in the stage game. The set of strategies for player 1 is denoted Σ_1 .

A strategy for player 2 at stage t is a map

$$\sigma_{2,t}: H_{2,t-1} \rightarrow S_2.$$

We let $\Sigma_{2,t}$ be the set of such strategies and let $\Sigma_2 = \prod_{t \geq 1} \Sigma_{2,t}$ denote the set of sequences of such strategies.

A history h_t of length t is an element of $\Omega \times (A_1 \times A_2 \times Z_1 \times Z_2)^t$ describing player 1's type and all actions and signals to the players up to stage t . A strategy

²The model includes the case in which z_1 reveals a_1 , in which case player 1 has perfect recall.

profile $\sigma = (\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ induces a probability distribution P_σ over the set of plays $H_\infty = \Omega \times (A_1 \times A_2 \times Z_1 \times Z_2)^{\mathbb{N}}$ endowed with the product σ -algebra. We let $a_t = (a_{1,t}, a_{2,t})$ represent the pair of actions chosen at stage t and let $z_t = (z_{1,t}, z_{2,t})$ denote the pair of signals received at stage t . Given $\omega \in \Omega$, $P_{\omega, \sigma}(\cdot) = P_\sigma(\cdot | \omega)$ represents the distribution over plays conditional on player 1 being of type ω .

Player 1’s discount factor is $0 < \delta < 1$ and the expected discounted payoff to player 1 of a normal type is

$$\pi_1(\sigma) = \mathbf{E}_{P_{\omega, \sigma}}(1 - \delta) \sum_{t \geq 1} \delta^{t-1} u_1(a_t).$$

2.2. ε -Entropy-Confirming Best Responses

The relative entropy (see Cover and Thomas (1991) for more on information theory) between two probability distributions P and Q over a finite set X such that the support of Q contains that of P is

$$d(P \parallel Q) = \mathbf{E}_P \ln \frac{P(x)}{Q(x)} = \sum_{x \in X} P(x) \ln \frac{P(x)}{Q(x)}.$$

The uniform norm distance between two probability measures over X is

$$\|P - Q\| = \max_{x \in X} |P(x) - Q(x)|.$$

Pinsker’s (1964) inequality implies that $d(P \parallel Q) \geq 2\|P - Q\|^2$.

Assume that player 1’s strategy in the stage game is α_1 and player 2 plays a best response α_2 to his belief α'_1 on player 1’s strategy. Under (α_1, α_2) , the distribution of signals to player 2 is $q^2(\cdot | \alpha_1, \alpha_2)$, while it is $q^2(\cdot | \alpha'_1, \alpha_2)$ under (α'_1, α_2) . Player 2’s prediction error on his signals, measured by the relative entropy, is $d(q^2(\cdot | \alpha_1, \alpha_2) \parallel q^2(\cdot | \alpha'_1, \alpha_2))$.

We define ε -entropy-confirming best responses in a similar way as the ε -confirming best responses of Fudenberg and Levine (1992). The main difference is that player 2’s prediction errors are measured using the relative entropy instead of the uniform norm. Another, albeit minor, difference is that we do not require player 2’s strategies to be non-weakly dominated.

DEFINITION 1: $\alpha_2 \in S_2$ is an ε -entropy-confirming best response to $\alpha_1 \in S_1$ if there exists α'_1 such that the following conditions hold:

- α_2 is a best response to α'_1 .
- $d(q^2(\cdot | \alpha_1, \alpha_2) \parallel q^2(\cdot | \alpha'_1, \alpha_2)) \leq \varepsilon$.

$B_{\alpha_1}^d(\varepsilon)$ denotes the set of ε -entropy-confirming best responses to α_1 .

Pinsker’s inequality implies that every non-weakly dominated ε -entropy-confirming best response is a $\sqrt{\varepsilon/2}$ -confirming best response. The sets of 0-entropy-confirming best responses and 0-confirming best responses coincide.

When player 1’s strategy is α_1 and player 2 plays an ε -entropy-confirming best response to it, the payoff to player 1 is bounded below by³

$$\underline{v}_{\alpha_1}(\varepsilon) = \min_{\alpha_2 \in B_{\alpha_1}^d(\varepsilon)} u_1(\alpha_1, \alpha_2).$$

The maximum payoff to player 1 if player 2 plays an ε -entropy-confirming best response to player 1’s strategy is

$$\bar{v}(\varepsilon) = \max\{u_1(\alpha_1, \alpha_2), \alpha_1 \in S_1, \alpha_2 \in B_{\alpha_1}^d(\varepsilon)\}.$$

The supremum of all convex functions below \underline{v}_{α_1} is denoted \underline{w}_{α_1} , and \bar{w} represents the infimum of all concave functions above \bar{v} . Practical computations and estimations of the maps \underline{v}_{α_1} , \underline{w}_{α_1} , \bar{v} , and \bar{w} are presented in Section 3.

Our main result below is proven in Appendix A.

THEOREM 1: *If σ is a Nash equilibrium, then*

$$\sup_{\hat{\omega}} \underline{w}_{\hat{\omega}}(- (1 - \delta) \ln \mu(\hat{\omega})) \leq \pi_1(\sigma) \leq \bar{w}(- (1 - \delta) \ln \mu(\tilde{\omega})).$$

Define the *lower Stackelberg payoff* as $\underline{v}^* = \sup_{\alpha_1 \in S_1} \underline{v}_{\alpha_1}(0)$. It is the least payoff player 1 obtains when choosing α_1 and player 2 plays a best response while predicting accurately the distribution of his own signals.

The *upper Stackelberg payoff* is $\bar{v}^* = \bar{v}(0)$. It is the maximal payoff to player 1 when player 2 plays a best response while predicting accurately the distribution of his own signals.

Let $\underline{N}(\delta)$ and $\bar{N}(\delta)$ be the infimum and the supremum, respectively, of Nash equilibrium payoffs to player 1 with discount factor δ . As a corollary to Theorem 1, we obtain the following version of Fudenberg and Levine’s (1992) results. The proof can be found in Appendix B.

COROLLARY 1—Fudenberg and Levine (1992):

$$\limsup_{\delta \rightarrow 1} \bar{N}(\delta) \leq \bar{v}^*.$$

If $\hat{\Omega}$ is dense in S_1 , then

$$\liminf_{\delta \rightarrow 1} \underline{N}(\delta) \geq \underline{v}^*.$$

³It is a consequence of Lemma 7 in Appendix B that the min and max below are well defined.

3. ESTIMATING THE VALUE OF REPUTATION FUNCTIONS

The bounds provided by Theorem 1 rely on knowledge of the functions $w_{\hat{\omega}}$ and \bar{w} . We show how these functions can be estimated, and that the bounds are sharper when monitoring is more accurate.

3.1. Lower Bounds

We first consider the benchmark case of perfect monitoring; then we consider games with imperfect monitoring.

3.1.1. Games With Perfect Monitoring

We assume here that player 2 has perfect monitoring in the stage game, that is, $Z_2 = A$, and $q^2(a|a) = 1$ for every a . As in Fudenberg and Levine (1989, pp. 765–766), we deduce from the upper hemicontinuity of the best response correspondence of player 2 that there exists d^* such that $B_{\hat{\omega}}^d(d) = B_{\hat{\omega}}^d(0)$ for every $d < d^*$. It follows that $v_{\hat{\omega}}(d) = v_{\hat{\omega}}(0)$ for $d < d^*$ and we bound $v_{\hat{\omega}}(d)$ by $\underline{u} = \min_a u_1(a)$ for $d \geq d^*$. It is then verified piecewise that $v_{\hat{\omega}}(d)$ is bounded below by the linear map $\frac{d}{d^*}\underline{u} + (1 - \frac{d}{d^*})v_{\hat{\omega}}(0)$ and it follows that $w_{\hat{\omega}}(d)$ admits the same lower bound. Theorem 1 shows that

$$(1) \quad \underline{N}(\delta) \geq -(1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*} \underline{u} + \left(1 + (1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*} \right) v_{\hat{\omega}}(0).$$

In case $\hat{\omega}$ is a pure strategy, Fudenberg and Levine (1989) proved the result

$$(2) \quad \underline{N}(\delta) \geq (1 - \delta^{\ln \mu(\hat{\omega}) / \ln \mu^*}) \underline{u} + \delta^{\ln \mu(\hat{\omega}) / \ln \mu^*} v_{\hat{\omega}}(0),$$

where μ^* is such that for every α'_1 , every best response to $\mu^* \hat{\omega} + (1 - \mu^*) \alpha'_1$ is also a best response to $\hat{\omega}$.

To compare the bounds (1) and (2), we first relate the highest possible value of d^* with the smallest value for μ^* . Observe that since $\hat{\omega}$ is a pure strategy, $d(\hat{\omega} \parallel \alpha_1) = -\ln \alpha_1(\hat{\omega}) < -\ln(\mu^*)$ if and only if $\alpha_1(\hat{\omega}) > \mu^*$. Hence the relationship $d^* = -\ln(\mu^*)$. In the interesting case where $\mu(\hat{\omega}) < \mu^*$, $\delta^{\ln \mu(\hat{\omega}) / \ln \mu^*} > 1 + (1 - \delta) \ln \mu(\hat{\omega}) / d^*$ and we conclude, not too surprisingly, that (2) is tighter than (1). The difference between the two bounds can be understood as coming from the fact that (2) is based on an exact study of player 2's beliefs path per path, whereas (1) comes from a study of beliefs on average. Note, however, that both bounds converge to $v_{\hat{\omega}}(0)$ at the same rate when $\delta \rightarrow 1$, since

$$\lim_{\delta \rightarrow 1} \frac{\delta^{\ln \mu(\hat{\omega}) / \ln \mu^*} - 1}{(1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*}} = 1.$$

Note that, unlike (1), (2) does not extend to the case in which $\hat{\omega}$ is a mixed strategy. In that case, Fudenberg and Levine (1992, Theorem 3.1) (see also

Proposition 15.4.1 in Mailath and Samuelson (2006)) instead proved that for every ε , there exists K such that for every δ ,

$$(3) \quad \underline{N}(\delta) \geq (1 - (1 - \varepsilon)\delta^K)\underline{u} + (1 - \varepsilon)\delta^K \underline{v}_{\hat{\omega}}(0).$$

To obtain the convergence of the lower bound to $\underline{v}_{\hat{\omega}}(0)$ requires first taking ε small enough, thus fixing K , then taking δ close enough to 1. Since K gets larger and larger as $\varepsilon \rightarrow 0$, (3) does not yield a rate of convergence of the lower bound to $\underline{v}_{\hat{\omega}}(0)$.

On the other hand, the bound (1) is explicit and provides a rate of convergence of $\underline{N}(\delta)$ as $\delta \rightarrow 1$ for mixed commitment types that is no slower than the rate implied by (2) for pure types.

EXAMPLE 1: To illustrate the bound (1) when commitment types are in mixed strategies, consider a perfect monitoring version of the quality game of Fudenberg and Levine (1989, Example 4) (see also Mailath and Samuelson (2006, Example 15.4.2)) in which player 1 can produce a good of high (H) or low (L) quality, and player 2 can decide to buy the good (b) or not (n):

	n	b
H	0, 0	1, 1
L	0, 0	3, -1

We identify player 1’s strategies with the probability they assign to H and assume the existence of a type $\hat{\omega} > \frac{1}{2}$ such that $\mu(\hat{\omega}) > 0$. Let $d^* = d(\hat{\omega} \parallel \frac{1}{2}) = \ln(2) + \hat{\omega} \ln(\hat{\omega}) + (1 - \hat{\omega}) \ln(1 - \hat{\omega}) > 0$. The unique best response to $\alpha'_1 > \frac{1}{2}$ is b , so that $B_{\hat{\omega}}^d(d) = \{b\}$ for every $d < d^*$. We have $\underline{v}_{\hat{\omega}}(0) = 3 - 2\hat{\omega}$ and $\underline{u} = 0$, and (1) becomes

$$\underline{N}(\delta) \geq \left(1 + (1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*}\right) (3 - 2\hat{\omega}).$$

Observe that the closer $\hat{\omega}$ is to $\frac{1}{2}$, the closer $3 - 2\hat{\omega}$ is to the Stackelberg payoff of 2, but the lower d^* is, hence the lower $1 + (1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*}$ is. This illustrates the trade-off according to which commitment types closer to $\frac{1}{2}$ yield higher payoffs in the long run, but may require a longer time to establish the proper reputation.

3.1.2. Games With Identifiable Actions

We say that player 2 identifies player 1’s actions whenever $\alpha_1 \neq \alpha'_1$ implies $q^2(\cdot | \alpha_1, \alpha_2) \neq q^2(\cdot | \alpha'_1, \alpha_2)$ for every α_2 . Let d_0^* be such that $d(\hat{\omega} \parallel \alpha_1) < d_0^*$ implies that every best response to α_1 is also a best response to $\hat{\omega}$. By continuity of the relative entropy and using the identifiability assumption, we can find d^*

such that $d(q^2(\cdot|\hat{\omega}, \alpha_2) \parallel q^2(\cdot|\alpha_1, \alpha_2)) < d^*$ implies $d(\hat{\omega} \parallel \alpha_1) < d_0^*$. We now have $B_{\hat{\omega}}^d(d) = B_{\hat{\omega}}^d(0)$ for every $d < d^*$ and we deduce as in Section 3.1.1 that

$$(4) \quad \underline{N}(\delta) \geq -(1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*} \underline{u} + \left(1 + (1 - \delta) \frac{\ln \mu(\hat{\omega})}{d^*} \right) v_{\hat{\omega}}(0).$$

Hence, similar lower bounds on player 1’s lowest equilibrium payoff, linear in $(1 - \delta)\mu(\hat{\omega})$, obtain under full monitoring and under the weaker assumption of identifiability. The formula remains the same, but the parameter d^* in (4) needs to be adjusted for the quality of monitoring. We show in Section 3.3 that poorer monitoring leads to weaker bounds.

3.1.3. Games Without Identifiable Actions

In games with perfect monitoring, and more generally in games with identifiable actions, the map $v_{\hat{\omega}}$ is constant and equal to $v_{\hat{\omega}}(0)$ in a neighborhood of 0. This is not the case when not all actions of player 2 allow for statistical identification of player 1’s actions as in the following example.

EXAMPLE 2: We take up the classical *entry game* from [Kreps and Wilson \(1982\)](#) and [Milgrom and Roberts \(1982\)](#). The stage game is an extensive-form game in which player 2 (the entrant) first decides whether to enter (E) a market or not (N), and following an entry of player 2, player 1 (the incumbent) decides whether to fight (F) or accommodate (A). The payoff matrix is the entry game

	E	N	
F	$-1, b - 1$	$a, 0$	
A	$0, b$	$a, 0$,

where $a > 1$ and $1 > b > 0$. It is natural to assume that each player 2 observes the outcomes of the previous interactions, namely, whether each past player 2 decided to enter or not, and whether player 1 decided to fight or accommodate each entry of player 2. Thus, $Z^2 = \{N, F, A\}$. Consider the pure strategy commitment type $\hat{\omega} = F$ and let q represent the probability that player 2’s mixed strategy assigns to E . Values $q \in (0, 1)$ are best responses only to $\alpha'_1 = bF + (1 - b)A$. Then, for $q \in (0, 1)$, $q^2(\cdot|\hat{\omega}, \alpha_2) = qF + (1 - q)N$ and $q^2(\cdot|\alpha'_1, \alpha_2) = qbF + q(1 - b)A + (1 - q)F$, and $d(q^2(\cdot|\hat{\omega}, \alpha_2) \parallel q^2(\cdot|\alpha'_1, \alpha_2)) = -q \ln(b)$. It follows that $B_F^d(\varepsilon) = \{q, q \leq -\frac{\varepsilon}{\ln b}\}$ and

$$w_F(\varepsilon) = v_F(\varepsilon) = u_1\left(F, -\frac{\varepsilon}{\ln b}\right) = a + (a + 1) \frac{\varepsilon}{\ln b}.$$

Now using Theorem 1, we obtain the bound

$$\underline{N}(\delta) \geq a - \frac{(a + 1)(1 - \delta)}{\ln b} \ln(\mu(F)).$$

In this example again, the rate of convergence of the lower bound on equilibrium payoffs is linear in $(1 - \delta)\mu(\hat{\omega})$.

3.2. Upper Bounds

3.2.1. Games With Perfect Monitoring

Let $U = \max_{a_1, a'_1, a_2} |u_1(a_1, a_2) - u_1(a'_1, a_2)|$. The following lemma is proven in Appendix C.

LEMMA 1: *We have*

$$\bar{v}(\varepsilon) \leq \bar{w}(\varepsilon) \leq \bar{v}^* + U\sqrt{\frac{\varepsilon}{2}}.$$

The upper bound from Theorem 1 and Lemma 1 together imply

$$(5) \quad \bar{N}(\delta) \leq \bar{v}^* + \frac{U}{\sqrt{2}}\sqrt{-(1 - \delta)\ln \mu(\tilde{\omega})}.$$

Fudenberg and Levine (1992) showed that for every ε , there exists K such that for every δ ,

$$(6) \quad \bar{N}(\delta) \leq (1 - (1 - \varepsilon)\delta^K) \max_a u_1(a) + (1 - \varepsilon)\delta^K \bar{v}_\omega(0).$$

Note that (6) provides an upper bound on $\bar{N}(\delta)$ that converges to \bar{v}^* as $\delta \rightarrow 1$. However, unlike from (5), no explicit rate of convergence can be derived from (6).

3.2.2. Games With Identifiable Actions

We take up games in which player 2 identifies player 1's actions, as in Section 3.1.2. We rely on the following lemma, proven in Appendix C.

LEMMA 2: *If player 2 identifies player 1's actions, then there exists a constant M such that for every $\alpha_1, \alpha'_1 \in S_1$ and $\alpha_2 \in S_2$,*

$$d(q^2(\cdot|\alpha_1, \alpha_2) \parallel q^2(\cdot|\alpha'_1, \alpha_2)) \geq M\|\alpha_1 - \alpha'_1\|^2.$$

Following the same line of proof as Lemma 1, we obtain that for M defined by Lemma 2,

$$\bar{v}(\varepsilon) \leq \bar{w}(\varepsilon) \leq \bar{v}^* + U\sqrt{M\varepsilon}.$$

Hence, the upper bound from Theorem 1 implies

$$\bar{N}(\delta) \leq \bar{v}^* + U\sqrt{-M(1 - \delta)\ln \mu(\tilde{\omega})}.$$

Thus, in games with identifiable actions as well as in games with full monitoring, we obtain an upper bound on $\bar{N}(\delta)$ that converges to \bar{v}^* when $\delta \rightarrow 1$ at a speed which is constant times $\sqrt{1 - \delta}$.

3.3. The Effect of the Quality of Monitoring

Consider an alternative monitoring structure q' with set of signals S'_2 for player 2. Assume that monitoring of player 2 is poorer under q' than under q in the sense that there exists $Q: S_2 \rightarrow \Delta(S'_2)$ such that $q'(s'_2|a_1, a_2) = \sum_{s_2} q'(s'_2|a_1, a_2)Q(s_2)(s'_2)$ for every a_1, a_2 . This is the case when player 2's signals under q' can be reproduced, using Q , from player 2's signals under q . Let $(q \otimes Q)^2(\cdot|\alpha_1, \alpha_2)$ and $(q \otimes Q)^2(\cdot|\alpha'_1, \alpha_2)$ be the probability distributions over $S_2 \times S'_2$ induced by $q^2(\cdot|\alpha_1, \alpha_2)$ and $q'^2(\cdot|\alpha'_1, \alpha_2)$, respectively, and Q . The chain rule of relative entropies (see Appendix A.1) implies that for every $\alpha_1, \alpha'_1, \alpha_2$,

$$\begin{aligned} & d(q^2(\cdot|\alpha_1, \alpha_2) \parallel q'^2(\cdot|\alpha'_1, \alpha_2)) \\ &= d((q \otimes Q)^2(\cdot|\alpha_1, \alpha_2) \parallel (q \otimes Q)^2(\cdot|\alpha'_1, \alpha_2)) \\ &\leq d(q^2(\cdot|\alpha_1, \alpha_2) \parallel q'^2(\cdot|\alpha'_1, \alpha_2)), \end{aligned}$$

where the inequality uses the fact that the marginals of $(q \otimes Q)^2(\cdot|\alpha_1, \alpha_2)$ and $(q \otimes Q)^2(\cdot|\alpha'_1, \alpha_2)$ over S'_2 are $q^2(\cdot|\alpha_1, \alpha_2)$ and $q'^2(\cdot|\alpha'_1, \alpha_2)$. This implies that the sets of ε -entropy best responses under the monitoring structure q are subsets of their counterparts under the monitoring structure q' . Let $\underline{v}'_{\alpha_1}(\varepsilon)$ and $\bar{v}'(\varepsilon)$ be the equivalents of $\underline{v}_{\alpha_1}(\varepsilon)$ and $\bar{v}(\varepsilon)$, defined under the monitoring structure q' instead of q . We obtain the following theorem.

THEOREM 2: *If monitoring is poorer under q' than under q , then for every α_1 and ε , $\underline{v}_{\alpha_1}(\varepsilon) \geq \underline{v}'_{\alpha_1}(\varepsilon)$ and $\bar{v}(\varepsilon) \leq \bar{v}'(\varepsilon)$.*

The theorem shows that the bounds equilibrium payoffs following from Theorem 1 are tighter under more informative structures for player 2 than for less informative structures.

APPENDIX A: PROOF OF THEOREM 1

A.1. The Chain Rule for Relative Entropies

Our proof relies on the fundamental property of relative entropies called the chain rule. Let P and Q be two distributions over a finite product set $X \times Y$ with marginals P_X and Q_X on X . Let $P_Y(\cdot|x)$ and $Q_Y(\cdot|x)$ be P 's and Q 's conditional probabilities on Y given $x \in X$. According to the chain rule of relative entropies,

$$d(P \parallel Q) = d(P_X \parallel Q_X) + \mathbf{E}_{P_X} d(P_Y(\cdot|x) \parallel Q_Y(\cdot|x)).$$

An implication of the chain rule is the following bound on the relative entropy between two distributions under some “grain of truth.”

LEMMA 3: *Assume $Q = \varepsilon P + (1 - \varepsilon)P'$ for some $\varepsilon > 0$ and P' . Then*

$$d(P \parallel Q) \leq -\ln \varepsilon.$$

PROOF: Consider the following experiment. (i) Draw a Bernoulli random variable X with probability of success ε . (ii) Conditional on $X = 1$, draw Y according to P ; conditional on $X = 0$, draw X according to P' . Let Q^0 be the corresponding distribution of $X \times Y$. Let P^0 be defined as Q^0 , except that the probability of success is 1. Denote by P_X^0 and P_Y^0 (resp., Q_X^0 and Q_Y^0) X and Y 's distributions under P^0 (resp., Q^0). Then according to the chain rule conditioning on X ,

$$d(P^0 \parallel Q^0) = d(P_X^0 \parallel Q_X^0) = -\ln(\varepsilon).$$

Using the chain rule conditioning on Y gives

$$d(P^0 \parallel Q^0) \geq d(P_Y^0 \parallel Q_Y^0) = d(P \parallel Q). \qquad \text{Q.E.D.}$$

A.2. Bound on Player 2's Prediction Errors

To obtain bounds on equilibrium payoffs, we first bound the expected discounted distance between player 2's prediction on his own signals and the true distribution of these signals given player 1's type.

For $n \geq 1$, consider the marginal $P_\sigma^{2,n}$ of P_σ over the sequences $H_{2,n}$ of signals of player 2, and for $\omega \in \Omega$, let $P_{\omega,\sigma}^{2,n}$ be the marginal of $P_{\omega,\sigma}$ over $H_{2,n}$.

LEMMA 4: *For $\omega \in \Omega$,*

$$d(P_{\omega,\sigma}^{2,n} \parallel P_\sigma^{2,n}) \leq -\ln \mu(\omega).$$

PROOF: Decompose $P_\sigma^{2,n} = \mu(\omega)P_{\omega,\sigma}^{2,n} + (1 - \mu(\omega))P'$, where P' is $P_\sigma^{2,n}$ conditioned on player 1 not being of type ω . The result follows from Lemma 3. Q.E.D.

Under P_σ , player 2's beliefs on the next stage's signals following $h_{2,t}$ are given by

$$p_\sigma^2(h_{2,t})(z_{2,t+1}) = P_\sigma(z_{2,t+1} | h_{2,t}).$$

We compare $p_\sigma^2(h_{2,t})$ with player 2's belief if player 2 knew that player 1 is of type ω , given by

$$p_{\omega,\sigma}^2(h_{2,t})(z_{2,t+1}) = P_{\omega,\sigma}(z_{2,t+1} | h_{2,t}).$$

The expected discounted sum of relative entropies between these two predictions is

$$d_{\omega,\sigma}^\delta = (1 - \delta) \sum_{t \geq 1} \delta^{t-1} \mathbf{E}_{P_{\omega,\sigma}} d(p_{\omega,\sigma}^2(h_{2,t}) \parallel p_\sigma^2(h_{2,t})).$$

Lemma 5 provides an upper bound on the expected total discounted prediction error of player 2.

LEMMA 5: *For every $\omega \in \Omega$,*

$$d_{\omega,\sigma}^\delta \leq -(1 - \delta) \ln \mu(\omega).$$

PROOF: From an iterated application of the chain rule,

$$d(P_{\omega,\sigma}^{2,n} \parallel P_\sigma^{2,n}) = \sum_{t=1}^n \mathbf{E}_{P_{\omega,\sigma}} d(p_{\omega,\sigma}^2(h_{2,t}) \parallel p_\sigma^2(h_{2,t})).$$

Using the decomposition of the discounted average as a convex combination of arithmetic averages (see, e.g., equation (1) in Lehrer and Sorin (1992)) and Lemma 4, we get

$$\begin{aligned} d_{\omega,\sigma}^\delta &= (1 - \delta)^2 \sum_{n \geq 1} \delta^{n-1} \sum_{t=1}^n \mathbf{E}_{P_{\omega,\sigma}} d(p_{\omega,\sigma}^2(h_{2,t}) \parallel p_\sigma^2(h_{2,t})) \\ &= (1 - \delta)^2 \sum_{n \geq 1} \delta^{n-1} d(P_{\omega,\sigma}^{2,n} \parallel P_\sigma^{2,n}) \\ &\leq (1 - \delta)^2 \sum_{n \geq 1} \delta^{n-1} (-\ln \mu(\omega)) \\ &= -(1 - \delta) \ln \mu(\omega). \end{aligned} \qquad \text{Q.E.D.}$$

A.3. Proof of the Lower Bounds on Equilibrium Payoffs

Consider a commitment type $\hat{\omega} \in \hat{\Omega}$ and let $\sigma_{1,\hat{\omega}} \in \Sigma_1$ be player 1’s strategy in which the normal type follows $\hat{\omega}$, given by $\sigma_{1,\hat{\omega}}(\tilde{\omega}, h_{1,t}) = \hat{\omega}$ for every $h_{1,t}$. The next lemma provides a lower bound on the payoff to player 1 of normal type playing $\sigma_{1,\hat{\omega}} \in \Sigma_1$ and facing an equilibrium strategy of player 2. It implies the lower bound of equilibrium payoffs of Theorem 1.

LEMMA 6: *If (σ_1, σ_2) is a Nash equilibrium, then for every $\hat{\omega} \in \hat{\Omega}$,*

$$\pi_1(\sigma_{1,\hat{\omega}}, \sigma_2) \geq \underline{w}_{\hat{\omega}}(-(1 - \delta) \ln \mu(\hat{\omega})).$$

PROOF: Let $\sigma' = (\sigma_{1,\hat{\omega}}, \sigma_2)$. Note that under σ' , and conditional on $h_{2,t} \in H_{2,t}$, the expected payoff to player 1 at stage $t + 1$ is $u_1(\hat{\omega}, \sigma_2(h_{2,t}))$. Since $\sigma_2(h_{2,t})$ is a $d(p_{\hat{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t}))$ -entropy-confirming best response to $\hat{\omega}$, then $u_1(\hat{\omega}, \sigma_2(h_{2,t})) \geq v_{\hat{\omega}}(d(p_{\hat{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t})))$.

Now we bound player 1's expected discounted payoff. Using $P_{\hat{\omega},\sigma'} = P_{\hat{\omega},\sigma}$, then $v_{\hat{\omega}} \geq w_{\hat{\omega}}$, and applying Jensen's inequality to the convex map $w_{\hat{\omega}}$, we obtain

$$\begin{aligned} \pi_1(\sigma') &= (1 - \delta) \sum_{t \geq 1} \delta^{t-1} \mathbf{E}_{P_{\hat{\omega},\sigma'}} u_1(\hat{\omega}, \sigma_2(h_{2,t})) \\ &\geq (1 - \delta) \sum_{t \geq 1} \delta^{t-1} \mathbf{E}_{P_{\hat{\omega},\sigma}} v_{\hat{\omega}}(d(p_{\hat{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t}))) \\ &\geq (1 - \delta) \sum_{t \geq 1} \delta^{t-1} \mathbf{E}_{P_{\hat{\omega},\sigma}} w_{\hat{\omega}}(d(p_{\hat{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t}))) \\ &\geq w_{\hat{\omega}} \left((1 - \delta) \sum_{t \geq 1} \delta^{t-1} \mathbf{E}_{P_{\hat{\omega},\sigma}} d(p_{\hat{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t})) \right) \\ &= w_{\hat{\omega}}(d_{\hat{\omega},\sigma}^{\delta}) \end{aligned}$$

and the result follows from Lemma 5, since $w_{\hat{\omega}}$ is nonincreasing. Q.E.D.

A.4. Proof of the Upper Bounds on Equilibrium Payoffs

Conditional on player 1 being of type $\tilde{\omega}$ and on history $h_{2,t}$ of player 2, the average strategy of player 1 at stage $t + 1$ is

$$\tau_1(h_{2,t}) = \sum_{h_{1,t}} P_{\tilde{\omega},\sigma}(h_{1,t}|h_{2,t}) \sigma_1(\tilde{\omega}, h_{1,t}).$$

Since $\sigma_2(h_{2,t})$ is a $d(p_{\tilde{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t}))$ -entropy-confirming best response to $\tau_1(h_{2,t})$, the payoff $u_1(\tau_1(h_{2,t}), \sigma_2(h_{2,t}))$ to player 1 of type $\tilde{\omega}$ at stage $t + 1$ conditional on $h_{2,t}$ is no more than $\bar{v}(d(p_{\tilde{\omega},\sigma}^2(h_{2,t}) \parallel p_{\sigma}^2(h_{2,t})))$.

As in the proof of Lemma 6, we deduce that $\pi_1(\sigma) \leq \bar{w}(d_{\tilde{\omega},\sigma}^{\delta})$, and the result follows from Lemma 5.

APPENDIX B: PROOF OF COROLLARY 1

LEMMA 7: *The map $(\alpha_1, \varepsilon) \mapsto B_{\alpha_1}^d(\varepsilon)$ is upper hemicontinuous.*

PROOF: Consider sequences $(\alpha_{1,n})_n \rightarrow \alpha_1$, $(\varepsilon_n)_n \rightarrow \varepsilon$, and $(\alpha_{2,n})_n \rightarrow \alpha_2$ with $\alpha_{2,n} \in B_{\alpha_1}^d(\varepsilon_n)$. Let $\alpha'_{1,n}$ be such that $\alpha_{2,n}$ is a best response to it and

$$d(q^2(\cdot|\alpha_{1,n}, \alpha_{2,n}) \parallel q^2(\cdot|\alpha'_{1,n}, \alpha_{2,n})) \leq \varepsilon_n.$$

Extract a subsequence $(\alpha'_{1,m(n)})_n$ of $(\alpha'_{1,n})_n$ converging to some α'_1 . From the lower semicontinuity of $d(\cdot \| \cdot)$,

$$d(q^2(\cdot|\alpha_1, \alpha_2) \| q^2(\cdot|\alpha'_1, \alpha_2)) \leq \varepsilon,$$

and the upper hemicontinuity of the best-response correspondence implies that α_2 is a best response to α'_1 . Hence, $\alpha_2 \in B^d_{\alpha_1}(\varepsilon)$. *Q.E.D.*

LEMMA 8: (i) \underline{v}_{α_1} and \underline{w}_{α_1} for every α_1 and $\sup_{\alpha_1 \in \hat{\Omega}} \underline{w}_{\alpha_1}$ are nonincreasing, lower semicontinuous, and continuous at 0, and $\underline{w}_{\alpha_1}(0) = \underline{v}_{\alpha_1}(0)$.

(ii) \bar{v}, \bar{w} are nondecreasing, upper semicontinuous, and continuous at 0, and $\bar{w}(0) = \bar{v}(0)$.

PROOF: We prove (i). The correspondence $B^d_{\alpha_1}(\cdot)$ being monotone and upper hemicontinuous, \underline{v}_{α_1} is nonincreasing and lower semicontinuous. These properties carry from \underline{v}_{α_1} to \underline{w}_{α_1} for every α_1 and to $\sup_{\alpha_1 \in \hat{\Omega}} \underline{w}_{\alpha_1}$, and they imply continuity at 0 for all these maps. $\underline{w}_{\alpha_1}(0) = \underline{v}_{\alpha_1}(0)$ is implied by the continuity at 0 of \underline{v}_{α_1} . Point (ii) is obtained by similar arguments. *Q.E.D.*

PROOF OF COROLLARY 1: From Theorem 1, $\limsup_{\delta \rightarrow 1} \bar{N}(\delta) \leq \bar{w}(\varepsilon)$ for every $\varepsilon > 0$, hence $\limsup_{\delta \rightarrow 1} \bar{N}(\delta) \leq \bar{w}(0) = \bar{v}(0) = \bar{v}^*$.

From Theorem 1, $\liminf_{\delta \rightarrow 1} \underline{N}(\delta) \geq \sup_{\alpha_1 \in \hat{\Omega}} \underline{w}_{\alpha_1}(\varepsilon)$ for $\varepsilon > 0$, hence $\liminf_{\delta \rightarrow 1} \underline{N}(\delta) \geq \sup_{\alpha_1 \in \hat{\Omega}} \underline{w}_{\alpha_1}(0) = \sup_{\alpha_1 \in \hat{\Omega}} \underline{v}_{\alpha_1}(0)$. By hemicontinuity of $\alpha_1 \mapsto B^d_{\alpha_1}(0)$, $\alpha_1 \mapsto \underline{v}_{\alpha_1}(0)$ is lower semicontinuous, and since $\hat{\Omega}$ is dense in S_1 , $\sup_{\alpha_1 \in \hat{\Omega}} \underline{v}_{\alpha_1}(0) = \sup_{\alpha_1 \in S_1} \underline{v}_{\alpha_1}(0) = \underline{v}^*$. *Q.E.D.*

APPENDIX C: PROOFS FROM SECTION 3

PROOF OF LEMMA 1: Since monitoring is perfect, $\alpha_2 \in B^d_{\alpha_1}(\varepsilon)$ implies that α_2 is a best response to $\alpha'_1 \in S_1$ such that $d(\alpha_1 \| \alpha'_1) \leq \varepsilon$. Hence

$$\begin{aligned} \bar{v}(\varepsilon) &= \max_{\alpha_1 \in S_1, \alpha_2 \in B^d_{\alpha_1}(\varepsilon)} \pi(\alpha_1, \alpha_2) \\ &\leq \max_{\alpha'_1 \in S_1, \alpha_2 \in B^d_{\alpha'_1}(0)} \pi(\alpha'_1, \alpha_2) \\ &\quad + \max_{\alpha_2 \in S_2, d(\alpha_1 \| \alpha'_1) \leq \varepsilon} |\pi_1(\alpha_1, \alpha_2) - \pi_1(\alpha'_1, \alpha_2)|. \end{aligned}$$

Note that $\max_{\alpha'_1 \in S_1, \alpha_2 \in B^d_{\alpha'_1}(0)} \pi(\alpha'_1, \alpha_2) = \bar{v}(0) = \bar{v}^*$, while Pinsker's inequality implies

$$\begin{aligned} &\max_{\alpha_2 \in S_2, d(\alpha_1 \| \alpha'_1) \leq \varepsilon} |\pi_1(\alpha_1, \alpha_2) - \pi_1(\alpha'_1, \alpha_2)| \\ &\leq \max_{\alpha_2 \in S_2, \|\alpha_1 - \alpha'_1\| \leq \sqrt{\varepsilon/2}} |\pi_1(\alpha_1, \alpha_2) - \pi_1(\alpha'_1, \alpha_2)| \end{aligned}$$

$$\begin{aligned} &\leq \max_{\|\alpha_1 - \alpha'_1\| \leq \sqrt{\varepsilon/2}} U \|\alpha_1 - \alpha'_1\| \\ &\leq U \sqrt{\frac{\varepsilon}{2}}. \end{aligned}$$

Since the map $\varepsilon \mapsto \bar{v}^* + U\sqrt{\frac{\varepsilon}{2}}$ is concave and lies above \bar{v} , it also lies above \bar{w} . *Q.E.D.*

PROOF OF LEMMA 2: From Pinsker’s inequality,

$$d(q^2(\cdot|\alpha_1, \alpha_2) \parallel q^2(\cdot|\alpha'_1, \alpha_2)) \geq 2\|q^2(\cdot|\alpha_1, \alpha_2) - q^2(\cdot|\alpha'_1, \alpha_2)\|^2.$$

To complete the proof, it is enough to show the existence of M' such that for every $\alpha_1, \alpha'_1, \alpha_2$,

$$\|q^2(\cdot|\alpha_1, \alpha_2) - q^2(\cdot|\alpha'_1, \alpha_2)\| \geq M' \|\alpha_1 - \alpha'_1\|,$$

which can be rewritten as

$$\begin{aligned} &\sum_{s_2} \left| \sum_{a_1, a_2} \alpha_2(a_2) (\alpha_1(a_1) - \alpha'_1(a_1)) q(s_2|a_1, a_2) \right| \\ &\geq M' \sum_{a_1} |\alpha_1(a_1) - \alpha'_1(a_1)|. \end{aligned}$$

Let

$$B = \left\{ \beta_1 : S_2 \rightarrow \mathbb{R}, \sum_{a_1} \beta_1(a_1) = 0, \sum_{a_1} |\beta_1(a_1)| = 1 \right\}.$$

The identifiability assumption shows that

$$\sum_{s_2} \left| \sum_{a_1, a_2} \alpha_2(a_2) \sum_{a_1} \beta_1(a_1) q(s_2|a_1, a_2) \right| > 0$$

for $\beta_1 \in B, \alpha_2 \in S_2$. Since $B \times A_1$ is compact, this implies the existence of M' such that

$$\sum_{s_2} \left| \sum_{a_1, a_2} \alpha_2(a_2) \sum_{a_1} \beta_1(a_1) q(s_2|a_1, a_2) \right| > M'$$

for every $\beta_2 \in B, \alpha_2 \in S_2$, hence the result.

Q.E.D.

REFERENCES

- COVER, T. M., AND J. A. THOMAS (1991): *Elements of Information Theory*. Wiley Series in Telecommunications. New York: Wiley. [1629]
- FUDENBERG, D., AND D. K. LEVINE (1989): "Reputation and Equilibrium Selection in Games With a Patient Player," *Econometrica*, 57, 759–778. [1627,1631,1632]
- (1992): "Maintaining a Reputation When Strategies Are Imperfectly Observed," *Review of Economic Studies*, 59, 561–579. [1627,1629-1631,1634]
- KREPS, D., AND R. WILSON (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279. [1633]
- LEHRER, E., AND S. SORIN (1992): "A Uniform Tauberian Theorem in Dynamic Programming," *Mathematics of Operations Research*, 17, 303–307. [1637]
- MAILATH, G., AND L. SAMUELSON (2006): *Repeated Games and Reputations: Long-Run Relationships*. New York: Oxford University Press. [1627,1632]
- MILGROM, P., AND J. ROBERTS (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280–312. [1633]
- PINSKER, M. (1964): *Information and Information Stability of Random Variables and Processes*. Holden-Day Series in Time Series Analysis. San Francisco: Holden Day. [1629]

Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris, France and London School of Economics, London, U.K.; ogossner@gmail.com.

Manuscript received June, 2010; final revision received December, 2010.