SIGNALLING GAMES AND INCENTIVE DOMINANCE*

Gonzalo Olcina & Amparo Urbano**

WP-AD 95-21

---

* We wish to thank financial aid from the Valencian Institute of Economic Research (IVIE).

** G. Olcina & A. Urbano: University of Valencia and I.V.I.E.

Editor: **Instituto Valenciano de Investigaciones Económicas, S.A.**
Primera Edición Septiembre 1995.
ISBN: 84-482-1123-5
Depósito Legal: V-3784-1995
Impreso por Copisteria Sanchis, S.L.,
Quart, 121-bajo, 46008-Valencia.

SIGNALLING GAMES AND INCENTIVE DOMINANCE

Gonzalo Olcina & Amparo Urbano

ABSTRACT

We present a new refinement for signalling games: the Introspective Equilibrium. It is based on both a procedure for beliefs formation -called Incentive Dominance- and a global consistence requirement, closely related to that of the Undefeated equilibrium of mailath, Okuno-Fujiwara and Postlewaite. The incentive Dominance criterion captures the principle of forward induction through explicitly modelling the players thought process when forming preliminary beliefs. The main idea is that they should exploit the information contained in the best reply structure about the incentives of the different types of a a rational Bayesian sender. Our criterion subsumes very intuitive ones as equilibrium dominance and divinity. The Introspective Equilibrium asks for an unambigous explanation of any deviation from a given equilibrium. This means that the explanation should be unique, part of a sequential equilibrium and achievable from the preliminary beliefs defined by the Incentive Dominance Criterion.

Keywords: Forward Induction; Incentive Dominance; Global Consistency.

1-.INTRODUCTION

Signalling games are very important because they represent a stylized model of a recurrent situation in economics: a transaction with private information on one side where costly signalling activities are possible.

The main issue in signalling games is that of "inference", that is, to construct conditional beliefs about the likelihood of different private information after seeing a particular signal from the informed party. The equilibrium refinements' literature for these games has tried to restrict beliefs off the equilibrium path, proposing distinct formulations of the so called principle of Forward Induction (FI).

Roughly speaking, a Sequential Equilibrium(SE) satisfies FI if it is consistent with deductions based on players' rational behavior in the past. The refinements' literature concretes this idea looking for a rational explanation of the deviation from an expected equilibrium. But the formalization of FI has proved to be elusive till the present. And, it is also quite well-known that there are some important logical difficulties in the foundations of this literature.

This is the so-called Stiglitz critique. The refinements' analysis is typically "local", that is, the analist takes a given sequential equilibrium and an unexpected message and tries to restrict beliefs off the equilibrium path in some reasonable way. Sometimes this yields different beliefs that the ones which support the given equilibrium. But, why do we stop the reasoning process at that point ? As Stiglitz pointed out, sometimes, if you follow the disequilibrium dynamics caused by the proposed refinement logic, you might get back to the original equilibrium. Hence, is or is not stable this equilibrium ?

In this paper we present a new equilibrium refinement for signalling games. We call it the *Introspective Equilibrium.* We adopt the conventional viewpoint of assuming the behavioral axiom consisting on replacing the equilibrium path by its expected payoff and analyzing one possible deviation at a time. In this setting our refinement captures the idea of FI by proposing a static solution concept that is the expression of the rest point of the (not explicitly modelled) players' thought process.

In particular, the two main ideas are, on one hand, that players should "exploit" the information contained in the best reply structure about the incentives of the different types of a rational Bayesian informed player. Thus, we establish a relationship between the assignment of conditional probabilities to the different types of informed player and the range of situations (responses of the uninformed player) for which deviating is better than playing the given equilibrium. This relationship reflects the principle of Insufficient Reason. We call this procedure for beliefs formation the *Incentive Dominance Criterion* (IDC). A good feature of this criterion is that it subsumes very popular refinements criteria as equilibrium dominance and divinity. A byproduct of our analysis is that we obtain a new characterization for these criteria in terms of an explicit model of beliefs formation. Nevertheless, if we define a refinement based on this criterion, it can be shown that it is not implied by strategic stability (Kohlberg and Mertens, 1986). The notion of Incentive Dominance captures an aspect of FI not captured by stability.

On the other hand, from the preliminary beliefs constructed in this way, players will engage in an introspective process until they achieve, if it is possible, a rational *-consistent-* explanation of the hypothetical deviation from the given equilibrium outcome.

The idea of consistent beliefs and actions is not new, i.e. Perfect Sequential Equilibrium (PSE, Grossman and Perry, 1986) and PSE* (Van Damme, 1987). However, this idea has not been included in a complete "story" about how people arrive at such consistent explanation. Unlike these refinements that assume common knowledge of the relevant consistent action, our concept selects it through a -not modeled- reasoning process that has the beliefs derived by the IDC as its starting point.

But, the rational explanation of the deviation has to satisfy two more conditions. The first one is "global" consistency, that is, it has to be part of another alternative SE (if there is any). This requisite guarantees that the Stiglitz critique does not apply to our proposal.

The second one,is that the rational explanation has to be "unambigous". This is a basic requisite that some authors have pointed out before (see, for example, van Damme 1987), but it has not been incorporated yet to the refinement literature. One possibility to achieve unambiguity is to ask for uniqueness of the rational explanation of the deviation. But, we think that this is too demanding because what is really important in signalling games is that there is no confussion about the subset of types which is likely to deviate.

We can think of the refinement literature as being divided in two categories. On one hand, the first category focuses on different proposals about reasonable restrictions on off the equilibrium path beliefs (for instance, the Intuitive Criterion, Divinity,..., just to mention the most popular ones). On the other hand, the other category stresses the requisite of consistent beliefs or actions (for instance, Perfect Sequential Equilibrium -PSE-, PSE*, ...). Both categories lack of a global consistency property and, in this sense, the Stiglitz critique can be applied to both of them. However, recent proposals, such as the Undefeated Equilibrium of Mailath et al., for example, incorporate the above property as their main motivation.

Our solution concept combines the first two categories and also asks for global consistency. Players form preliminary beliefs that satisfy reasonable properties but, they will follow the reasoning process from these beliefs until they find, if it exists, a rest point. This rest point has to take the form of a consistent belief-action pair and has to be part of a SE.

Instead of modelling this quite complex dynamics we propose a static notion that captures the properties of these rest points. In this sense, our proposal relates to the underlying eductive dynamic approach as the concept of Evolutive Stable Strategy (Maynard Smith) does to the Replicator Dynamics in an evolutive context.

The plan of this paper is as follows. In Section 2 we review some notation and define the auxiliary game that results from analyzing one deviation at a time. In Section 3 we derive the criterion of Incentive Dominance. In developing the argument we also obtain a new characterization of divinity. Section 4 presents and discusses the new refinement. Section 5 applies our concept to the well known Spence's job market model. Section 6 is devoted to analyze the relationships among our concept and other refinements and provides concluding remarks.

2-. THE MODEL AND THE AUXILIARY GAME

We limit attention to simple signalling games. In these games, one player, the Sender or player 1, receives private information. We refer to this information as the Sender's type, **t** ; **t** is drawn from a finite set T according to a probability distribution π over T. We assume that π is common knowledge and that $\pi(t) > 0$ for all $t \in T$. After the Sender learns his type, he sends a signal m from a finite set M to the other player, the Receiver or player 2. This one responds to the Sender's signal m by choosing an action, a, from a finite set of responses A. The players have von Neumann-Morgenstern utility functions defined over type, signal and action. The sender's payoff function is denoted u(t,m,a) and the Receiver's payoff function is denoted v(t,m,a) ; we extend these functions to the set of all mixed strategies by linearity and use u(.) and v(.) to refer to these extensions.

We will work with the set of best-replies of the Receiver. Let μ be a probability distribution over T, i.e. $\mu \in \Delta(T)$, and let

$$BR(\mu,m)= \operatorname*{Arg\,max}_{a\in A} \sum_{t\in T} v(t,m,a)\ \mu(t).$$

If the receiver thinks that $\mu(t)$ is the probability that the sender is type t given the signal m, then $BR(\mu,m)$ is the set of best replies to m. Let $BR(m)= \bigcup_{\mu\in\Delta(T)} BR(\mu,m)$ be the set of pure best replies of the receiver after signal m.We write $MBR(\mu,m)$ for the set of mixed best replies corresponding to $BR(\mu,m)$, and we denote co(X) for the convex hull of the set X.

A *sequential equilibrium* of a signalling game Γ consists of a behavior strategy for the sender, denoted by $p(m|t)$, which specifies the probability that the sender of type t sends the signal $m \in M$; a behavior strategy for the receiver, denoted by $q(a|m)$, which specifies the probability that the receiver takes the action $a \in A$ in response to the signal m; and assessments, denoted by $\mu(t|m)$, such that $\mu(.|m)$ is a probability distribution over T for each $m \in M$. A triple (p,q,μ) is a sequential equilibrium of a signalling game if and only if p is a best response to q ($p(m'|t) > 0$, only if m' maximizes $u(t, m, q(.|m))$ over all $m \in M$); the receiver responds optimally to his assessment ($q(.|m) \in MBR(\mu(.|m),m)$ for all $m \in M$); and the assessments are consistent with the equilibrium strategy of the sender and the prior whenever possible (if $\sum_{t'} \pi(t')p(m|t') > 0$, then $\mu(t|m)=\left[\pi(t)p(m|t)\right] / \left[\sum_{t'}\pi(t')p(m|t')\right]$). We will write SE($\Gamma$) for the set of SE of a signalling game Γ.

Fix a sequential equilibrium $e = (p,q,\mu)$ of a given signalling game. Let u*(e,t) be the equilibrium payoff for the sender of type t, and let m be an unsent signal in this equilibrium. Define a new game, called the *auxiliary* game, and denote it by G(e,m),where the sender's types have only two signals m* and m. If they send m*, they obtain the equilibrium payoff u*(e,t); if they send m, the receiver plays as in the original signalling game.

Define,

$$D(t,m)=\left\{\alpha \in co(BR(m)):\ u(t,m,\alpha)\geq u^*(e,t)\right\},$$

thus D(t,m) is the set of the sender's conjectures over the pure best responses of the receiver for which sending signal m is a best reply for type t in the auxiliary game G(e,m).

Also, let

$$D^0(t,m)= \left\{\alpha \in co(BR(m)) \colon u(t,m,\alpha)= u^*(e,t)\right\} ,$$

be the corresponding similar set for which m is a weak best response for t in G. Obviously,

$$D(t,m) \backslash D^0(t,m)= \left\{\alpha \in co(BR(m)) \colon u(t,m,\alpha) > u^*(e,t)\right\}.$$

3-. A CRITERION FOR PRELIMINARY BELIEFS FORMATION

This section relies heavily in Olcina and Urbano (1994), and has been included for completeness. Assume, as it is conventional, that the players expect a particular Sequential Equilibrium. Before actual play, they will analyze introspectively how they should interpret a deviation m from the expected equilibrium play. Therefore, provided they analyze one deviation at a time, this thought experiment is equivalent to reasoning about how to play in the Auxiliary game.

Player 2 has to think about rational motives of the different types of the sender to deviate and send signal m. In particular, he has a very valuable information : he knows the best-reply structure of G(e,m). This structure reflects the global preliminary incentives of Bayesian rational players, i.e. players that optimize against subjective conjectures. Therefore, he will use this best-reply structure in order to form subjective probabilities about type t choosing m. The main idea is that this probability has to be proportional to the range of situations for which sending m is better for type t than getting the equilibrium payoff. From this probability and the exogenous prior π the receiver can construct reasonable conditional beliefs $\mu(t|m)$. Let us be more specific.

The receiver will analyze separately the incentives to send signal m of each type t of the sender. He knows that the sender is "Bayesian". Therefore, whatever is his type, he will have a conjecture over the pure best responses of player 2. Let α be this conjecture. Obviously, $\alpha \in co(BR(m))$.

But player 2 knows that the sender is also a rational player. That is, he will play a best reply to his conjecture α. Obviously, the sender's best reply depends on his type t. Let us call $b_t(\alpha)$, type t's best reply. In particular,

$$b_t(\alpha) = \begin{cases} m & \text{if } u(t,m,\alpha) > u^*(t). \\ m^* & \text{if } u(t,m,\alpha) < u^*(t). \end{cases}$$

Obviously, when $u(t,m,\alpha) = u^*(t)$, type t will be indifferent between sending signal m or signal m*.

The receiver does not know the subjective beliefs α, but, as he is a Bayesian, he will assign subjective probabilities to the possible values of α. Let f_2 be this "second order" beliefs. Therefore, $f_2(x)$ is the probability that the receiver assigns to "1 believes that α is equal to x".

From this f_2 and the fact that the sender, whatever is his type, is Bayesian rational, the receiver can compute the probability that type t chooses m as :

$$\int_{\{x\,:\,u(t,m,\alpha)\,\geq\,u^*(t)\}} f_2(x)dx$$

(Recall that $\{x : u(t,m,\alpha) \geq u^*(t)\} = D(t,m)$.)

But the receiver, as a Bayesian player, should assign initially a probability to this event. Let us denote it by $\lambda(t,m)$. A minimal consistency condition implies :

$$\lambda(t,m) = \int_{D(t,m)} f_2(\alpha)d\alpha$$

Now it is clear that how does f_2 look like is determinant in order to construct the probabilities $\lambda(.,m)$.

Divinity and Equilibrium Dominance.

In principle there is no reason for the receiver to believe that 1 has one conjecture or another. Hence, let us do the weakest assumption and suppose that f_2 is positive in co(BR(m)).

Now, assume there are types t and t' such that $D(t,m) \subset D(t',m)\backslash D^0(t',m)$. Then, it is evident that whatever f_2 looks like, $\lambda(t',m) \geq \lambda(t,m)$.

Recall that $\lambda(t,m)$ is the probability of type t sending m. From this and the "exogenous prior" information $\pi(t)$, the receiver can compute $\mu(t|m)$, that is, the conditional probability of being t who sends signal m.

Therefore, from $\lambda(t',m) \geq \lambda(t,m)$ we conclude that $\mu(t',m)/\mu(t,m) \geq \pi(t')/\pi(t)$. In other words, t should not be more likely to send m than is t'. Readers familiar with the refinement literature would have recognized that we have obtained in a different way the criterion for beliefs restriction called Divinity (Banks and Sobel, 1987). To be more precise, a minor variation known as Co-divinity. Notice that in those cases in which co(BR(m))=MBR(m), then both divinity and co-divinity coincide.

Therefore we obtain a new characterization of Co-divinity (Divinity) in terms of an explicit model on beliefs formation. This characterization highlights the weak assumptions which are behind this concept.

Notice also that if $D(t,m) = \emptyset$, then $\lambda(t,m) = 0$ and, obviously, $\mu(t|m) = 0$. This case is equivalent to the criterion of Equilibrium Dominance (Cho and Kreps, 1987).

The general case : the incentive dominance criterion.

What can we say if there is no "inclusion" relation between the sets D(t,m) ? Nothing at all if we do not make more specific assumptions on the beliefs f_2. Given the existing strategic uncertainty and the fact that players are trying to figure out some kind of "preliminary theory", it seems natural to assume that all conjectures over the receiver's pure best replies are "equally likely". That is, f_2 is uniformly distributed on co(BR(m)).

But then this implies that $\lambda(t,m)$ is exactly the Lebesgue measure of the set D(t,m). Thus,

$$\lambda(t,m) = \int_{D(t,m)} f_2(\alpha)d\alpha = D^l(t,m),$$

where $D^l(t,m)$ is the Lebesgue measure of the set $D(t,m)$.

The sets $D(t,m)$ are the "stability sets" of signal m for the different types of sender. Therefore, we claim that their Lebesgue measure express the theoretical probability of type t finding himself in a situation in which m is his best response to the action of the receiver in $G(m)$.

Now, the receiver should combine this information with his prior π. One possibility is to obtain a "point belief", i.e. $\hat{\mu} \in \Delta(T)$ such that $\hat{\mu}(t) = c\pi(t)D^l(t,m)$, for all $t \in T$ and some $c > 0$.

Instead of that, we only impose a weaker and more natural way of "updating". Namely, suppose $D^l(t,m) > 0$ and $D^l(t',m) > 0$, if $D^l(t,m) \geq D^l(t',m)$, then $\mu(t)/\mu(t') \geq \pi(t)/\pi(t')$. In words, if sending m is better than not deviating from the equilibrium for type t for more conjectures over the pure best replies of the receiver than it is for type t', then the receiver should not raise the probability that the sender is t' relative to the probability that the sender is t.

Notice also that whenever $D(t,m) = \varnothing$, then $D^l(t,m) = 0$. Therefore, in this case $\lambda(t,m) = 0$ and it seems natural to require that $\mu(t) = 0$.

Definition 1: The set of beliefs that satisfy the Incentive Dominance Criterion, denoted by $\hat{\beta}(m,e)$ (or $\hat{\beta}$ when there is no confusion), is defined by the following two conditions:

(i) whenever $D^I(t,m) > 0$ and $D^I(t',m) > 0$, if $D^I(t,m) \geq D^I(t',m)$, then $\mu(t)/\mu(t') \geq \pi(t)/\pi(t')$, and

(ii) whenever $D^I(t,m) = 0$, then $\mu(t) = 0$.

Therefore, this set represents the preliminary beliefs after an unexpected message m, when players expect equilibrium e, and it reflects the "size" of the incentives to deviate of the different types.

We think in the Incentive Dominance Criterion (IDC) as the result of a boundedly rational model of preplay reasoning for forming preliminary beliefs. From this point of view, it relies in two simple assumptions. Namely, it is mutual knowledge that both players are Bayesian rational ("maximizers") and that they apply the Principle of Insufficient Reason to form the beliefs about the Sender's beliefs. This last assumption seems reasonable once we have eliminated the receiver's dominated actions and it is inspired by the heuristic justification of the risk dominance concept given by Harsanyi and Selten.

A first proposal of a sequential equilibrium refinement.

Given the IDC beliefs, player 2, after receiving message m , will be expected to play some action in $MBR(m,\hat{\beta}) := \bigcup_{\mu \in \hat{\beta}} MBR(m,\mu)$.

Now, we could define a refinement of sequential equilibrium, just by restricting beliefs to those allowed by the IDC.

Let us define $p(m) := \sum_t \pi(t)p(m|t)$.

Definition 2 : A sequential equilibrium (p,q,μ) of a signalling game Γ satisfies the Incentive Dominance Criterion (IDC) if for all m such that $p(m) = 0$, $q(.|m) \in MBR(m,\hat{\beta}(m,e))$ (provided $\hat{\beta}(m,e) \neq \emptyset$).

Next we prove that this refinement subsumes equilibrium dominance and co-divinity . Let $\beta^{CD}(m,e)$, be the set of beliefs that satisfies co-divinity, as defined above.

Definition 3: A Sequential Equilibrium $e=(p,q,\mu)$ of a signalling game is *co-divine* if for each m, such that $p(m)=0$, $q(.|m) \in MBR(m,\beta^{CD}(m,e))$.

Then, it can be shown:

PROPOSITION 3.1 : Every Sequential Equilibrium that satisfies the IDC is Co-divine.

Proof. It suffices to show that if a SE is not Co-divine, then it does not satisfy the IDC.

First, notice that $\hat{\beta} \subseteq \beta^{CD}$, i.e. $\mu \in \beta^{CD} \Rightarrow \mu \in \hat{\beta}$. This follows from the fact that if there exists a pair t and t' such that,

$$D(t,m) \subset D(t',m) \setminus D^0(t',m), \quad (*)$$

then all $\mu \in \beta^{CD}$ will satisfy the property:

$$\frac{\mu(t')}{\mu(t)} \geq \frac{\pi(t')}{\pi(t)} \qquad (**).$$

But (*) implies $D^1(t',m) > D^1(t,m)$, and then the IDC means that all $\mu \in \hat{\beta}$, also have to satisfy (**). If $D(t,m)=\emptyset$, co-divinity implies that $\mu(t)=0$. But in this case, $D^1(t,m)=0$, and the IDC also implies $\mu(t)=0$.

Next, $\hat{\beta} \subseteq \beta^{CD}$ implies that $MBR(m,\hat{\beta}) \subseteq MBR(m,\beta^{CD})$, because less conjectures justify less actions.

If a SE is not co-divine, there exists some unsent m, such that $q(.|m) \notin MBR(m,\beta^{CD})$. This implies that $q(.|m) \notin MBR(m,\hat{\beta})$. Hence this SE does not satisfy the IDC. ∎

As it is well known in the literature, co-divinity subsumes Equilibrium Dominance. Therefore from this fact and the above Proposition we have:

COROLLARY. Every Sequential Equilibrium that satisfies the IDC, satisfies Equilibrium Dominance.

Both Co-divinity and Equilibrium Dominance criteria on belief formation are usually applied iteratively, and so may be done with the IDC. Analogous results to those of the previuos Proposition and Corollary can be obtained for this case. However, since this is just a byproduct of our analysis, we do not extend on it.

4-. INTROSPECTIVE EQUILIBRIUM

The Incentive Dominance Criterion is just a criterion to form "preliminary" beliefs. In this sense, asking for the off the equilibrium path beliefs to satisfy the IDC is just some kind of "minimal" requirement. But, rational and intelligent players will continue the introspective process (from the beliefs defined by the IDC) until they find (if it exists) a "consistent" or "rational" explanation of the hypothetical deviation m. Let us illustrate this idea with the following example.

$$\begin{array}{ccc} & & \begin{array}{ccc} 1 & m & r \end{array} \\ \begin{bmatrix} 2,2 \\ 2,2 \\ 2,2 \end{bmatrix} & \begin{array}{c} t_1 \\ t_2 \\ t_3 \end{array} & \begin{bmatrix} 3,3 & 0,0 & 0,0 \\ 0,0 & 0,3 & 3,0 \\ 0,0 & 3,0 & 0,3 \end{bmatrix} \\ L & & R \end{array}$$

Game $\Gamma(1)$

$\pi(t_i) = 1/3, \; i = 1,2,3.$

Figure 1

In the game $\Gamma(1)$ there is a sequential equilibrium outcome in which all sender's types pool in signal L. Moreover, this outcome is divine and satisfies the never a weak best response property, but it does not belong to a stable component. It is easy to compute, given the symmetry of the game, that the only belief that satisfies the IDC is $\mu = (1/3,1/3,1/3)$. But, all mixed actions of 2 are best replies to μ. Therefore, this belief supports the SE where all t play L and 2 would answer to signal R with some mixture that assigns probability 2/3 at most to each of his pure actions.

But, the players should follow the introspective process from μ. In particular, notice that there is a unique candidate for "fixed point" of such a process, achievable from μ. Namely, action 1 is a best reply to μ, but if 2 plays l only t_1 would deviate and 2 should have a belief such that $\mu(t_1) = 1$. This, in turn, confirms that player 2 plays l. Therefore, the pair action-belief (l, $\mu(t_1) = 1$) satisfies a kind of fixed point condition of the introspective process. Notice that there is no other pair with this property.

In general, a *consistent* explanation (α,μ) is a pair action-belief, such that μ is the conditional distribution of π over types with incentive to deviate given that the receiver would play action α and α is a best reply to μ. The idea of consistent explanations is not new, i.e. Perfect Sequential Equilibrium ((PSE), Grossman and Perry, 1986), or PSE* (Van Damme, 1987). However, this idea has not been included in a complete "story" about how people arrive at such explanations.

Unlike the PSE*, our new refinement does not assume common knowledge of the relevant consistent explanation; instead of that it selects it through a reasoning process that has the beliefs derived by the IDC as its starting point. In other words, it is the static expression of the rest points of this reasoning dynamics.

There are two other important differences between what we have in mind and the PSE*concept : the consistent explanation has to be part of another SE in order to be inmune to the Stiglitz critique. This is the case in the previous example, game $\Gamma(1)$, where type t_1 sending R and the receiver responding with action l is part of a separating equilibrium.

The other difference is that we consider unavoidable,. in an eductive context, that a rational explanation has to be unambiguos. If there are several alternative equilibria, achievable from the IDC beliefs, in which a given disequilibrium message is sent, the receiver migth be uncertain about which one is the relevant consistent explanation and this uncertainty can, in turn, support the initial equilibrium. Consider, for example, the game $\Gamma(2)$ and the pooling equilibrium outcome in L.

$$\begin{array}{cccc} & & \begin{array}{ccc} l & m & r \end{array} & \Gamma(2) \\ \begin{bmatrix} 2,2 \\ 2,2 \\ 2,2 \end{bmatrix} & \begin{matrix} t_1 \\ t_2 \\ t_3 \end{matrix} & \begin{bmatrix} 3,3 & 0,0 & 0,0 \\ 0,0 & 3,3 & 0,0 \\ 0,0 & 0,0 & 3,3 \end{bmatrix} & \pi(t_i)=1/3, \quad i=1,2,3 \\ L & & R & \end{array}$$

Figure 2

Like in $\Gamma(1)$ the unique belief that satisfies the IDC is $\mu=(1/3,1/3,1/3)$. However, there are three possible consistent explanations, achievable from this belief, namely the action-belief pairs: $(l,\mu(t_1)=1)$, $(m,\mu(t_2)=1)$, and $(r,\mu(t_3)=1)$ and all of them are part of some alternative SE. Obviously, there is no reason to consider one pair more likely than others, so that the receiver cannot reduce his strategic uncertainty through introspection. Hence it seems very sensible to conclude that each type is equally likely, and this supports the pooling equilibrium outcome in L.

Notice that what is relevant to achieve unambiguity is not the uniqueness of the alternative equilibria. The important requisite is that, in all alternative equilibria achievable from the IDC beliefs, the set of types who send the message and prefer it to the given equilibrium payoff is the same, i.e. it is "unique". In the next section, we will ilustrate this point when we analyze the Spence's model.

We formalize next the above intuitive arguments.

Let e=(p,q,μ) be a SE of a signalling game Γ and m an unsent message in this equilibrium, i.e. p(m)=0. Define the following subset E(m,e) of SE of the game Γ (i.e. E(m,e)C SE(Γ)).

Definition 4: E(m,e) is the subset of SE(Γ) with elements e'=(p',q',μ'), such that p'(m) > 0, and

(i) $q'(.|m) \in MBR(m,\hat{\beta}(m,e))$,

(ii) $\forall$ t $\in$ K(e'): $u^*(t,e') \geq u^*(t,e)$, and $\exists$ t $\in$ K(e'):$u^*(t,e')>u^*(t,e)$, where $K(e')= \left\{t \in T: p'(t,m) >0)\right\}$, and

(iii) K(e')=K(e''), for $\forall$ e',e'' $\in$ E, e'$\neq$e''.

Therefore, E(m,e) is the set of rational explanation candidates for the unexpected deviation m in the SE, e. In other words, (1) it is the set of globally consistent explanations, that is, part of an alternative SE, (2) achievable from "reasonable" preliminary beliefs (the IDC beliefs), and (3) unambiguous in the sense of uniqueness of the set of types who might be interested in deviating.

Definition 5: A sequential equilibrium e=(p,q,μ), of a signalling game Γ is an *Introspective Equilibrium* if for each m, such that p(m)=0, $q(.|m)=q'(.|m)$ for some e'= (p',q',μ') $\in$ E(m,e), whenever E(m,e)$\neq \emptyset$.

Let us illustrate with an example how this refinement works. Consider the following game $\Gamma(3)$

$$\begin{array}{c} \\ \begin{bmatrix} 2,2 \\ 2,2 \end{bmatrix} \\ L \end{array} \begin{array}{c} \\ t_1 \\ t_2 \\ \\ \end{array} \begin{array}{c} \begin{array}{cc} l & r \end{array} \\ \begin{bmatrix} 3,3 & 0,2 \\ 0,0 & 3,2 \end{bmatrix} \\ R \end{array} \qquad \begin{array}{c} \Gamma(3) \\ \\ \\ \pi(t_i)=1/2, \quad i=1,2 \end{array}$$

Figure 3

The pooling equilibrium outcome in L is divine. It is supported by the belief $\mu(t_1|R)=2/3$. Let α be the probability that player 2 chooses l after R. Then the pooling L is supported by any $\alpha \in [1/3, 2/3]$.

Straightforward calculations yield that:

$D(t_1,R)=\left\{\alpha \ : \ \alpha \geq 2/3\right\}$, and $D(t_2,R)=\left\{\alpha \ : \ \alpha \leq 1/3\right\}$. Therefore, since the Lebesgue measure of these sets is the same, the unique belief consistent with the IDC is $\mu=1/2$. The only best reply to this belief is action r, hence type t_2 would deviate, and then the pooling in L is not an Introspective Equilibrium. Notice that in this game the set of SE where R is sent is not a singleton (and, in each SE is a different type who sends R), meanwhile the set E(R, pooling in L) has a unique element (namely, the separating SE where t_2 sends R). This illustrates how the IDC helps to select the relevant consistent explanation.

5-. THE SPENCE'S JOB MARKET SIGNALLING MODEL

In this section we will apply our solution concept to a simple version of the job market signalling model of Spence (1973) in which there are just two types of workers and where education is not productivity increasing. This model has been extensively analyzed and we will use it to illustrate how our equilibrium refinement works and which are the new insights that it offers.

Consider the situation described by the following rules :

i) Player 1 is a worker, coming out from a population where there are workers who (exogenously) have either high or low productivity. For simplicity, suppose high productivity means marginal productivity of 2 and low means marginal productivity of 1. It is common knowledge that the proportion of high productivity workers in the population is p. In game theoretical language, chance determines whether player 1 is of type $t = 1$ or $t = 2$, i.e. the set of types is $T = \{ 1,2 \}$.

ii) The worker, knowing t, selects an education level e, where $e \geq 0$.

iii) Simultaneously, two firms offer non-negative wages $w_1(e)$ and $w_2(e)$, respectively, based on the observed education level e.

iv) The worker selects an employer.

We assume that the cost of acquiring the education level e is [e/t] for a worker with marginal productivity (type) t. We also assume that all players are risk neutral. Hence, the payoff function of type t is $u_t = w - e/t$.

We will restrict attention to subgame perfect equilibria where the worker always chooses the firm offering the highest wage, and randomizes equally in case both firms offer the same wage. It is also quite well-known that the following is equivalent to stage iii) : we assume there is a single firm (player 2) offering a single wage w, based on the observed e and with payoff : $-(t - w)^2$.

Let us call μ for the posterior probability of t = 2 given an observed education level e, i.e. $\mu = \text{prob}\ (t = 2 | e)$. Therefore, in stage iii) two firms are competing a la Bertrand for a surplus of expected value $(1+\mu)$. The unique Nash-Bertrand equilibrium is for both firms to offer $w = 1 + \mu$. If we model this stage as a single player 2 with payoff function $-(t-w)^2$, then , it is clear that his best reply function for any belief μ, is $w=1 + \mu$. Notice also that any wage offer not in [1,2] is strictly dominated. Therefore, in what follows, we will only consider $w \in [1,2]$.

The set of sequential equilibria.

In this model there is a plethora not only of Nash equilibria, but also of SE. Let us characterize this latter set. First of all we describe the pure strategy SE.

There is a continuum of separating pure strategy SE, that is, equilibria in which the two types of workers choose different education levels e_1 and e_2. In particular, in all these SE, $e_1 = 0$, $e_2 \in [1,2]$ and the equilibrium wage function is $w(e_1) = 1$ and $w(e_2) = 2$. This equilibrium path is supported by the out of equilibrium wage offer $w(e) = 1$, for $e \neq e_1, e_2$, which is, in turn, supported on the disequilibrium beliefs that any education level different from the equilibrium ones is chosen by the low productivity worker.

There is also a continuum of pooling pure strategy equilibria in which both types of worker choose the same education level ê. Namely, $e_1 = e_2 = \hat{e}$, where $\hat{e} \in [0,p]$ and $w(\hat{e}) = 1+p$. Again $w(e) = 1$, for $e \neq \hat{e}$, supported on the belief $\mu = 0$.

In addition to all these pure strategy SE there are also hybrid equilibria in which some type randomizes among two education levels.

On one hand, there is a set of hybrid SE in which the high productivity type chooses an education level $e^* \in (p,1)$ and the low productivity type randomizes between e^* and $e_1 = 0$, with probabilities $(1 - q_1)$ and q_1 respectively. The equilibrium wage function is $w(e_1) = 1$ and $w(e^*) = 1 + \mu$, where, $\mu = p / [p + (1-p)(1-q_1)] > p$. Notice that $w(e^*) \in (1+p,2)$.

On the other hand, there is another continuum of hybrid equilibria in which type $t = 2$ randomizes. Namely, the low productivity type chooses an education level $e^* \in (0,p)$ and the high productivity type randomizes between e^* and e_2 with probability $(1-q_2)$ and q_2 respectively. Where $2-2p+e^* < e_2 < 2-e^*$, for each $e^* \in (0,p)$. The equilibrium wage function is $w(e_2)=2$ and $w(e^*)=1+\mu$, where $\mu=[p(1-q_2)]/[p(1-q_2)+(1-p)] < p$.

Notice that the set of education level with positive probability of being sent in some SE is the interval [0,2]. Therefore, in order to check if a SE is an IE we should disregard any education level greater than 2. Also note, that for most of the messages in the interval [0,2], when considered as off the equilibrium path messages in a given SE, there is more than one alternative SE where this education level is sent with positive probability. But this does not represent any problem because it is always the same set of types which send the message under consideration.

The set of introspective equilibria.

We obtain next, the set of *Introspective Equilibria* (IE) in this model as a corollary of a set of Propositions.

PROPOSITION 5.1. There is no separating Equilibrium in the Spence's model with $e_2>1$ which is an Introspective Equilibrium.

Proof. It is enough to show one education level e' different from e_2 and 0, for which the IE definition fails. Take the out of equilibrium message e'=1. Then,

$$D(t=1,e')=\left\{w \in [1,2]: w-1 \geq 1\right\}, \text{ and}$$

$$D(t=2,e')=\left\{w \in [1,2]: w-1/2 \geq 2-e_2/2 \right\}.$$

Therefore, in any of these separating equilibria, $D^1(t=1),e')=0$, and $D^1(t=2,e')\neq 0$, which in turn implies that the Incentive Dominance beliefs are $\mu=1$, and $w(e')=2$.

For $p \leq 1/2$, there is a unique SE where e' is sent by t=2; this is the Riley outcome and it is preferred by this type (its payoff is 1'5) to any of the equilibria under consideration.

For $p > 1/2$, there is also an hybrid SE in which t=2 sends e' and obtains a payoff of 1'5. Therefore, there is no ambiguity about which type is deviating to e'. ∎

PROPOSITION 5.2. There is no pooling equilibria in the Spence's model with $\hat{e} \in (0,p]$, which is an Introspective Equilibrium.

Proof. Take the out of equilibrium message e'=0. The SE where e' is sent with positive probability are: the pooling in e', all separating SE and the hybrid equilibria where t=1 randomizes.

These equilibria are supported by the ID beliefs. Namely,

$$D(t=1,\ e')=\left\{w \in [1,2]:\ w \geq 1+p-\hat{e}\right\},\ \text{and}$$
$$D(t=2,\ e')=\left\{w \in [1,2]:\ w\geq 1+p-\hat{e}/2\right\}.$$

It is easy to check that both sets are non-empty and furthermore, D(t=2,e')C D(t=1,e'). This implies that the set of ID beliefs is characterized by $\mu \leq p$.

But, in all the separating and hybrid equilibria the type that sends e', namely t=1, gets a payoff of 1, which is less or equal than the equilibrium payoff in any of the pooling equilibria under consideration.

Therefore the set E only contains the pooling in e' and both types obtain a higher payoff than the equilibria under consideration ($1+p > 1+p-\hat{e}$, and $1+p > 1+p-\hat{e}/2$, for $\hat{e} \in (0,p]$). ∎

PROPOSITION 5.3. For $p \geq 1/2$, the pooling equilibrium in the Spence's model, with $\hat{e}=0$, is an Introspective Equilibrium. For $p < 1/2$, this equilibrium is not an IE.

Proof. Let us check first of all that this pooling is not an IE if $p < 1/2$. Take the out of equilibrium message e'=1. There is a separating equilibrium where t=2 sends e' and gets a payoff of 1'5, which is greater than 1+p. Easy calculations show that D(t=1, e') is the empty set and D(t=2, e')is nonempty. Therefore, $\mu=1$ is the only ID belief which supports the separating equilibrium.

Next, we prove that for $p \geq 1/2$, the pooling in $\hat{e}=0$ is an IE. For any $e' \in (0,p]$, there exist pooling and hybrid SE in which both types pool in e'. But, in all of them both types obtain a lower payoff than in the candidate equilibrium (furthermore, the hybrid equilibria are not supported by ID beliefs). Then, the set E is empty.

For $e' \in (p,1)$, there are again hybrid SE in which t=1 randomizes between 0 and pooling in e' with t=2; and also there are hybrid SE in which t=2 randomizes between e' and pooling in some $\bar{e} \in (0,p)$ with t=1. But, again, both types obtain a lower payoff than in the pooling in e=0.

The same argument applies to any $e' \in [1,2]$. The IDC does not restrict the beliefs, but in any case, the types who send the message in the alternative equilibria obtain a lower payoff than $1+p$. Therefore the set E is empty. Note, that for $p=1/2$ and $e'=1$, type 2 gets 1'5 in the Riley outcome and then this type is indifferent to the pooling in $e=0$. ■

PROPOSITION 5.4. For $p \leq 1/2$, the separating equilibrium in the Spence's model, with $e_2=1$ (The Riley outcome), is an Introspective Equilibrium. For $p > 1/2$, this equilibrium is not an IE.

Proof. First of all, let us prove that for $p>1/2$ the Riley outcome is not an IE. Take, for example, the out of equilibrium message $e'=2p-1$. Notice that e' belongs to $(0,p)$. There is a pooling SE in e' where $t=2$ gets a payoff of: $1+p-[2p-1]/2$ =1'5, and $t=1$ gets $1+p-[2p-1]=2-p >1$. Therefore, type 1 is better off than under the Riley outcome and type 2 is indifferent. Notice, that there is also an hybrid with $t=2$ sending e' with positive probability but this type obtains a lower payoff.

Next, let us check the ID beliefs: $D(t=1, e')=\left\{w \in [1,2]: w \geq 2p\right\} \neq \emptyset$, and $D(t=2, e')=\left\{w \in [1,2]: w\geq 1+p\right\} \neq \emptyset$. Since $D(t=2, e')$ C $D(t=1,e')$,then $\mu \leq p$. Hence, the pooling is the unique element of the set E.

Now consider $p \leq 1/2$. For any $e' \in (0,p]$, any pooling and/or hybrid in e' yield a lower payoff to type 2. Therefore, it is not justified to pool in e'. A similar argument works for the cases in which $e' \in (p,1)$ and $e'(1,2]$: type 2 gets a lower payoff than in the Riley outcome.

Notice that when $p=1/2$, the pooling in $e=0$ yield the same payoff to type 2 and a better payoff to type 1; but, $e=0$ is not an out of equilibrium message in the Riley outcome. ■

The reader may easily check that none of the hybrid equilibria in the Spence's model is an IE.

COROLLARY. The set of Introspective Equilibria of the Spence's model is the following:

The pooling equilibrium in $\hat{e}=0$, if $p > 1/2$.

Both the pooling equilibrium in $\hat{e}=0$ and the separating equilibrium in which $t=2$ chooses $e=1$, if $p=1/2$.

The separating equilibrium in which $t=2$ chooses $e=1$, if $p < 1/2$.

Remarks:

An interesting feature, which we conjecture that can be generalized to more general monotonic signalling games, is that it is enough, in order to obtain the same results, with Codivine preliminary beliefs, instead of the IDC beliefs. That is, we can replace condition i) in Definition 4 by $q'(.|m) \in MBR(m,\beta^{CD}(m,e))$.

Thus, if someone feels uncomfortable with the uniformity assumption on the receiver's second order beliefs made in the construction of the IDC, we can see that the results carry on just assuming positive second order beliefs (which is the weakest assumption one can postulate).

It may be of interest to compare our solution of the Spence's model with the predictions of other solutions concepts. In particular, it is surprising that the set of IE coincides with the selection obtained applying Harsanyi and Selten's general theory of equilibrium selection (Van Damme and Güth, 1991). Both approaches obtain the SE which is best from the point of view of the high productivity worker.

It is very well known that strategic stability selects the Riley outcome for all $p \in (0,1)$. In fact, with two types it is enough with the Intuitive Criterion. Therefore, a conclussion of our analysis is that the Riley outcome is not immune to the "Stiglitz critique" when the proportion of high productivity workers in the population is high ($p > 1/2$).

A refinement which asks for consistent beliefs, as the Perfect Sequential Equilibrium, predicts the Riley outcome for $p < 1/2$, but fails to obtain a prediction for $p > 1/2$. In the latter case, the non existence problem is caused by the lack of a global consistency requirement.

As we will further comment in the next section, the set of IE in this model coincides with the set of a closely related refinement, called the Undefeated Equilibrium which also incorporates a global consistency condition.

6-. RELATIONSHIP WITH OTHER REFINEMENTS AND CONCLUDING REMARKS

The refinement concept closest to ours is that of *Undefeated Equilibrium* (Mailath et al.,1993). Roughly speaking , and quoting these authors : consider a proposed SE and a message that is not sent in the equilibrium. Suppose there is an alternative SE in which some non-empty set of types of player 1 choose the given message and that that set is precisely the set of types who prefer the alternative equilibrium to the proposed equilibrium. Then, they require that player 2's beliefs at that action in the original equilibrium to be consistent with this set. If they are not, we say the second equilibrium defeats the proposed one. Obviously, a SE is undefeated if there is no alternative SE that defeats it. More formally,

Definition 6 (Mailath, Okuno-Fujiwara and Postlewaite, 1993) : The (pure strategy) SE e' defeats the (pure strategy) SE e if there is a disequilibrium message m in e such that :

a) $K = \left\{ t \in T : p'(m,t) = 1 \right\} \neq \emptyset$.

b) $\forall\ t \in K,\ u^{*}(e',t) \geq u^{*}(e,t)$ and $\exists\ t \in K,\ u^{*}(e',t) > u^{*}(e,t)$.

c) $\exists\ t \in K$ such that, $\mu(t \mid m) \neq \left[\pi(t)\delta(t)\right] \,/\, \left[\sum_{t'} \pi(t')\delta(t')\right]$,

for any $\delta : T \to [0,1]$ satisfying $t' \in K$ and $u^{*}(e',t) > u^{*}(e,t) \Rightarrow \delta(t') = 1$ and $t' \notin K \Rightarrow \delta(t') = 0$.

A Sequential equilibrium is undefeated if there is no other sequential equilibrium that defeats it.

The concept of undefeated equilibrium does not consider mixed strategy equilibria. In this sense, it is weaker than the IE. But, apart from this fact, the right and interesting comparison is among undefeated and pure strategy IE. It is easy to define our solution concept only for pure strategy SE and to interpret it in terms of the "defeated " definition. In this context, the IE concept asks for more conditions in order to a SE being defeated, i.e. it would make less likely that an equilibrium is defeated. In particular, the SE which defeats another one has to be unambiguous in two senses : supported in the IDC beliefs and with respect to the set of types who might deviate. Therefore, the IE is weaker than the undefeated equilibrium. In the following proposition we prove formally this conjecture.

PROPOSITION 6.1. Consider a signalling game Γ, then generically, any undefeated equilibrium is an introspective equilibrium in pure strategies.

Proof. We show that if a SE e is not an IE, then it is defeated by another SE e'. We restrict attention to generic signalling games in which any type from the sets K(m,e') is better off deviating and in which the receiver's response at m in e'is unique.

If the SE $e = (p,q,\mu)$ is not an IE, then there is a message m such that $p(m) = 0$, $E(m,e) \neq \emptyset$ and for all $e' \in E(m,e)$, $q'(.|m) \neq q(.|m)$.

Take any $e' = (p',q',\mu') \in E(m,e)$. We show that e' defeats e. We have to verify that conditions a), b) and c) in Definition 6 are fulfilled.

Condition ii) in the Definition 4 (of the set E(m,e)) implies conditions a) and b). Moreover, notice that, as we restrict attention to pure strategy equilibria, $K(e') = \left\{t \in T : p'(t,m) = 1\right\}$ and $u^*(e',t) > u^*(e,t)$ for

$\forall$ t $\in$ K(e'). Therefore, the function δ is unique and updating π with δ yields a unique posterior.

Denote this posterior by π^k. We have to prove that $\mu \neq \pi^k$, which is condition c). Assume, to the contrary, that $\mu = \pi^k$. Notice that μ'(m,t) is the equilibrium belief in the SE e', derived using Bayes' formula from the equilibrium strategies. Therefore, $\mu' = \pi^k$. Given that the receiver's response is unique and $q'(.|m) \in MBR(\mu')$, $q(.|m) \in MBR(\mu)$, then we conclude that q = q'. But this result contradicts the assumption that the SE e is not an IE. ∎

Notice that in the Spence's model analyzed in the previous section the set of IE coincides with the set of undefeated equilibria, as it is easy to confirm. However, for general signalling games, as the above Proposition shows, IE is weaker, i.e. it gives weaker predictions in some games. This is caused by conditions i) and ii) in the definition of the sets E(m,e).

In essence, these two conditions ask for an unambiguous explanation of the deviation in two senses : it has to be achievable from some reasonable restricted beliefs and the set of types who might deviate in alternative SE has to be unique.

We believe that both requirements are very important in an eductive context, as we have explained along all the paper. The reader can complete his opinion taking a look to some examples. The game $\Gamma(2)$ in Figure 2 illustrates the importance of "unambiguity with respect to the set of types". In this game the pooling in L is defeated (by any of the separating equilibria), but, as we explained before, it is an IE.

The next game $\Gamma(4)$ ilustrates how the IDC beliefs matter.

$$\begin{array}{cccc} & & \begin{array}{ccc} l & m & r \end{array} & \\ \begin{bmatrix} 1'5\ ,0 \\ 1'5\ ,0 \end{bmatrix} & \begin{array}{c} t_1 \\ t_2 \end{array} & \begin{bmatrix} 2,3 & 0,2 & -1,0 \\ 1,0 & 2,2 & 0,3 \end{bmatrix} & \\ L & & R & \end{array}$$

$\Gamma(4)$

$\pi(t_i) = 1/2,\ i = 1,2.$

Figure 4

In this game there are two SE : Pooling in L, supported by player 2 planning to response r in case of receiving R, and a Separating SE where t_1 sends R (and 2 responds l) and t_2 sends L. In the pooling SE player 2 plans to play r after R because he believes that t_2 is more likely to deviate. It is easy to see that pooling in L is defeated by the separating SE. But, we will see that this pooling is an IE.

The reader can check easily that $D^I(t_2,R) > D^I(t_1,R)$, therefore, the IDC beliefs are those which assign more probability to type t_2, i.e. $\mu \leq 1/2$. But , the beliefs that support playing action l are $\mu \geq 2/3$. Therefore, the set E is empty and the pooling SE is an IE. In fact, it is supported by the belief that it is more likely that type t_2 deviates and these are very reasonable beliefs given the types incentives. Intuitively, one can think in the following "reasoning dynamics". Given the IDC beliefs, the receiver could plan to respond with action r, but then this action supports pooling in L. Alternatively, he might play m. Then, only type t_2 would send message R , justifying again the receiver's response r.

As the examples illustrate, there is an important difference in the "philosophy" of both concepts - the IE and the undefeated equilibrium. In particular, we have not in mind any "story" about "mistaken equilibria" and we do not see any clear heuristic behind the comparisons among SE of the undefeated concept definition. Instead of that, the story behind our proposal is an eductive one : the players reasoning process trying to achieve an explanation of the deviation. All the conditions that this explanation has to satisfy, including that it has to be part of another SE, have to be thought as necesary conditions for the explanation to be a rest point of the underlying, but not modelled, reasoning dynamics.

It is quite clear that the IE is not implied, nor does imply equilibrium dominance and/or divinity. For example, it is well-known that the unique SE outcome that does survive the Intuitive Criterion in the above Spence's model is the Riley outcome. As we have proved the set of IE in this model is quite different.

We have seen that a stable equilibrium may not be an IE. Therefore, we cannot apply the standard existence proof for refinements implied by stability. In fact, it is easy to construct examples where the set of IE is empty. This feature is shared with other refinements such as the PSE and PSE^*, or the undefeated equilibrium. The examples of non existence are very similar, so that we refer the reader to those of Mailath et al.

In our opinion the non existence of IE in some games shows the limitations of the refinement's analysis. This is typically a "local" analysis: we take as given one equilibrium and one deviation and look for a rational explanation of it. In this way, it is possible to fall in "cycles" in some particular games.

In any case, the IE exists in many important subclasses of signalling games; for example, the Spence's model of the previous section. By

Proposition 6.1., if the set of undefeated equilibria is non-empty, so it is the set of (pure strategy) IE. For example, this seems quite clear for the class of monotonic signalling games analyzed by Mailath et al. and many others. Anyway, it belongs to further research the precise characterization of classes of signalling games for which IE exists, classes for which it coincides with undefeated , etc...

To summarize, we have presented a new refinement for signalling games: the Introspective Equilibrium. Its main property is that it asks for a "consistent with introspection" explanation of any deviation from a given equilibrium. In particular, this means that this explanation is unambiguous, part of a sequential equilibrium and achievable from the preliminary beliefs defined by the Incentive Dominance criterion. This procedure for beliefs formation reflects the information contained in the best reply structure about the incentives of the different types of a rational Bayesian Sender.

A merit of this new concept is that it integrates two lines of research in the refinements literature. Namely, the one based on restrictions on beliefs and the other based on their consistence.

REFERENCES

Banks, J. S. and Sobel, J., Equilibrium Selection in Signalling Games., *Econometrica,* **55** (1987), 647-661.

Cho, I.K. and Kreps, D., Signalling Games and Stable Equilibria., *Quartely Journal of Economics*, **102** (1987), 179-221.

Donze, J., Globally Consistent Equilibria in Signaling Games., Mimeo. (1994)

Grossman, S.J. and Perry, M., Perfect Sequential Equilibrium, *Journal of Economic Theory*, **39** (1986), 120-154.

Kohlberg, E., Refinement of Nash Equilibrium: The Main Ideas. Mimeo, (1989) Harvard University.

Kohlberg, E. and Mertens, J-F., On the Strategic Stability of Equilibria, *Econometrica* **54** (1986), 1003-1037.

Mailath, G. J., Okuno-Fujiwara, M. and Postlewaite, A., Beliefs-Based Refinements in Signalling Games, *Journal of Economic Theory* **60** (1993), 241-276.

Olcina, G. and Urbano, A., A Note on Beliefs Formation in Signalling Games, *Economics Letters* **44** (1994), 55-59.

Sobel, J., Stole, L. and Zapater, I. , Fixed-Equilibrium Rationalizability in Signalling Games, *Journal of Economic Theory* **52** (1990), 304-331.

Van Damme, E. and Güth, W.,"Equilibrium Selection in the Spence Signaling Game." in Game Equilibrium Models. Vol. II, R. Selten (Ed.) Springer -Verlag, 1991.

Van Damme, E., "Stability and Perfection of Nash Equilibria". Springer-Verlag, 1987.

PUBLISHED ISSUES[*]

WP-AD 92-01	"Inspections in Models of Adverse Selection" I. Ortuño. May 1992.
WP-AD 92-02	"A Note on the Equal-Loss Principle for Bargaining Problems" C. Herrero, M.C. Marco. May 1992.
WP-AD 92-03	"Numerical Representation of Partial Orderings" C. Herrero, B. Subiza. July 1992.
WP-AD 92-04	"Differentiability of the Value Function in Stochastic Models" A.M. Gallego. July 1992.
WP-AD 92-05	"Individually Rational Equal Loss Principle for Bargaining Problems" C. Herrero, M.C. Marco. November 1992.
WP-AD 92-06	"On the Non-Cooperative Foundations of Cooperative Bargaining" L.C. Corchón, K. Ritzberger. November 1992.
WP-AD 92-07	"Maximal Elements of Non Necessarily Acyclic Binary Relations" J.E. Peris, B. Subiza. December 1992.
WP-AD 92-08	"Non-Bayesian Learning Under Imprecise Perceptions" F. Vega-Redondo. December 1992.
WP-AD 92-09	"Distribution of Income and Aggregation of Demand" F. Marhuenda. December 1992.
WP-AD 92-10	"Multilevel Evolution in Games" J. Canals, F. Vega-Redondo. December 1992.
WP-AD 93-01	"Introspection and Equilibrium Selection in 2x2 Matrix Games" G. Olcina, A. Urbano. May 1993.
WP-AD 93-02	"Credible Implementation" B. Chakravorti, L. Corchón, S. Wilkie. May 1993.
WP-AD 93-03	"A Characterization of the Extended Claim-Egalitarian Solution" M.C. Marco. May 1993.
WP-AD 93-04	"Industrial Dynamics, Path-Dependence and Technological Change" F. Vega-Redondo. July 1993.
WP-AD 93-05	"Shaping Long-Run Expectations in Problems of Coordination" F. Vega-Redondo. July 1993.
WP-AD 93-06	"On the Generic Impossibility of Truthful Behavior: A Simple Approach" C. Beviá, L.C. Corchón. July 1993.
WP-AD 93-07	"Cournot Oligopoly with 'Almost' Identical Convex Costs" N.S. Kukushkin. July 1993.

---

[*] Please contact IVIE's Publications Department to obtain a list of publications previous to 1992.

WP-AD 93-08 "Comparative Statics for Market Games: The Strong Concavity Case"
L.C. Corchón. July 1993.

WP-AD 93-09 "Numerical Representation of Acyclic Preferences"
B. Subiza. October 1993.

WP-AD 93-10 "Dual Approaches to Utility"
M. Browning. October 1993.

WP-AD 93-11 "On the Evolution of Cooperation in General Games of Common Interest"
F. Vega-Redondo. December 1993.

WP-AD 93-12 "Divisionalization in Markets with Heterogeneous Goods"
M. González-Maestre. December 1993.

WP-AD 93-13 "Endogenous Reference Points and the Adjusted Proportional Solution for Bargaining Problems with Claims"
C. Herrero. December 1993.

WP-AD 94-01 "Equal Split Guarantee Solution in Economies with Indivisible Goods Consistency and Population Monotonicity"
C. Beviá. March 1994.

WP-AD 94-02 "Expectations, Drift and Volatility in Evolutionary Games"
F. Vega-Redondo. March 1994.

WP-AD 94-03 "Expectations, Institutions and Growth"
F. Vega-Redondo. March 1994.

WP-AD 94-04 "A Demand Function for Pseudotransitive Preferences"
J.E. Peris, B. Subiza. March 1994.

WP-AD 94-05 "Fair Allocation in a General Model with Indivisible Goods"
C. Beviá. May 1994.

WP-AD 94-06 "Honesty Versus Progressiveness in Income Tax Enforcement Problems"
F. Marhuenda, I. Ortuño-Ortín. May 1994.

WP-AD 94-07 "Existence and Efficiency of Equilibrium in Economies with Increasing Returns to Scale: An Exposition"
A. Villar. May 1994.

WP-AD 94-08 "Stability of Mixed Equilibria in Interactions Between Two Populations"
A. Vasin. May 1994.

WP-AD 94-09 "Imperfectly Competitive Markets, Trade Unions and Inflation: Do Imperfectly Competitive Markets Transmit More Inflation Than Perfectly Competitive Ones? A Theoretical Appraisal"
L. Corchón. June 1994.

WP-AD 94-10 "On the Competitive Effects of Divisionalization"
L. Corchón, M. González-Maestre. June 1994.

WP-AD 94-11 "Efficient Solutions for Bargaining Problems with Claims"
M.C. Marco-Gil. June 1994.

WP-AD 94-12	"Existence and Optimality of Social Equilibrium with Many Convex and Nonconvex Firms" A. Villar. July 1994.
WP-AD 94-13	"Revealed Preference Axioms for Rational Choice on Nonfinite Sets" J.E. Peris, M.C. Sánchez, B. Subiza. July 1994.
WP-AD 94-14	"Market Learning and Price-Dispersion" M.D. Alepuz, A. Urbano. July 1994.
WP-AD 94-15	"Bargaining with Reference Points - Bargaining with Claims: Egalitarian Solutions Reexamined" C. Herrero. September 1994.
WP-AD 94-16	"The Importance of Fixed Costs in the Design of Trade Policies: An Exercise in the Theory of Second Best", L. Corchón, M. González-Maestre. September 1994.
WP-AD 94-17	"Computers, Productivity and Market Structure" L. Corchón, S. Wilkie. October 1994.
WP-AD 94-18	"Fiscal Policy Restrictions in a Monetary System: The Case of Spain" M.I. Escobedo, I. Mauleón. December 1994.
WP-AD 94-19	"Pareto Optimal Improvements for Sunspots: The Golden Rule as a Target for Stabilization" S.K. Chattopadhyay. December 1994.
WP-AD 95-01	"Cost Monotonic Mechanisms" M. Ginés, F. Marhuenda. March 1995.
WP-AD 95-02	"Implementation of the Walrasian Correspondence by Market Games" L. Corchón, S. Wilkie. March 1995.
WP-AD 95-03	"Terms-of-Trade and the Current Account: A Two-Country/Two-Sector Growth Model" M.D. Guilló. March 1995.
WP-AD 95-04	"Exchange-Proofness or Divorce-Proofness? Stability in One-Sided Matching Markets" J. Alcalde. March 1995.
WP-AD 95-05	"Implementation of Stable Solutions to Marriage Problems" J. Alcalde. March 1995.
WP-AD 95-06	"Capabilities and Utilities" C. Herrero. March 1995.
WP-AD 95-07	"Rational Choice on Nonfinite Sets by Means of Expansion-Contraction Axioms" M.C. Sánchez. March 1995.
WP-AD 95-08	"Veto in Fixed Agenda Social Choice Correspondences" M.C. Sánchez, J.E. Peris. March 1995.
WP-AD 95-09	"Temporary Equilibrium Dynamics with Bayesian Learning" S. Chatterji. March 1995.
WP-AD 95-10	"Existence of Maximal Elements in a Binary Relation Relaxing the Convexity Condition" J.V. Llinares. May 1995.

WP-AD 95-11	"Three Kinds of Utility Functions from the Measure Concept" J.E. Peris, B. Subiza. May 1995.
WP-AD 95-12	"Classical Equilibrium with Increasing Returns" A. Villar. May 1995.
WP-AD 95-13	"Bargaining with Claims in Economic Environments" C. Herrero. May 1995.
WP-AD 95-14	"The Theory of Implementation when the Planner is a Player" S. Baliga, L. Corchón, T. Sjöström. May 1995.
WP-AD 95-15	"Popular Support for Progressive Taxation" F. Marhuenda, I. Ortuño. May 1995.
WP-AD 95-16	"Expanded Version of Regret Theory: Experimental Test" R. Sirvent, J. Tomás. July 1995.
WP-AD 95-17	"Unified Treatment of the Problem of Existence of Maximal Elements in Binary Relations. A Characterization" J.V. Llinares. July 1995.
WP-AD 95-18	"A Note on Stability of Best Reply and Gradient Systems with Applications to Imperfectly Competitive Models" L.C. Corchón, A. Mas-Colell. July 1995.
WP-AD 95-19	"Redistribution and Individual Characteristics" I. Iturbe-Ormaetxe. September 1995.
WP-AD 95-20	"A Mechanism for Meta-Bargaining Problems" Mª. C. Marco, J. E. Peris, B. Subiza. September 1995.
WP-AD 95-21	"Signalling Games and Incentive Dominance" G. Olcina, A. Urbano. September 1995.