Selecting a pooling equilibrium in a signaling game with a bounded set of signals

Miguel Ángel Ropero

Department of Applied Economics

Faculty of Economics

University of Malaga

El Ejido, 6, 29071, Malaga, Spain

E-mail address: maropero@uma.es

Telephone number: +34 952 13 11 84

Fax number: +34 952 13 66 16

Abstract

In this paper, we study a general class of monotone signaling games, in which the support of the signal is limited or the cost of the signal is sufficiently low and as a result, there are multiple pooling equilibria. In those games, when we relax the usual single-crossing condition, the typical restrictions on the out-of-equilibrium beliefs suggested by previous literature cannot discard any of the equilibria obtained. For this reason, we develop a new refinement called the most profitable deviator, which will be useful to select a unique equilibrium in those games. Additionally, when the standard single-crossing condition is satisfied, our criterion also chooses a unique equilibrium, which is the same as that selected by previous literature.

Keywords: Most profitable deviator, out-of-equilibrium beliefs, sequential equilibrium, signaling game, single-crossing condition.

JEL Classification: C72, C73, C79.

1. Introduction

In a typical signaling game, an economic agent, the sender, has private information on some hidden characteristic and uses that information to make a decision, which is usually known as the signal or message sent by the sender to a second economic agent, the receiver. In order to simplify the theoretical setting, it is usually assumed that the hidden characteristic can take different values and each of these values will identify a "type" of sender. After observing the decision made by the sender, the receiver will form some beliefs about the type of sender against whom he or she is playing and given those beliefs, the receiver will take an action. Finally, both the sender and the receiver will get their pay-offs, which will depend on the sender´s type and on both players´ decisions. In most of these signaling games, there are multiple Perfect Bayesian equilibria and previous literature has proposed different criteria, which are also called refinements, in order to eliminate some of those equilibria.

In most signaling games studied in economics, only the separating equilibrium proposed by Riley (1979) survives the previous refinements (Cho and Kreps, 1987), but the selection of this equilibrium crucially depends on two assumptions. First, the signal takes a sufficiently large number of values. Second, the sender´s cost of the signal increases or decreases sufficiently with the sender´s type. In those signaling games in which any of these assumptions is not satisfied, multiple pooling equilibria will arise and previous refinements do not reduce the number of equilibria.

In this paper, we propose a binding refinement in such games. According to the criterion proposed, the receiver´s belief outside an equilibrium will be restricted using the following line of thought. If sending an out-of-equilibrium message might make a type of sender increase his or her pay-off by more than another type in comparison to what they get in equilibrium, the receiver should attach a greater probability to the first type of sender than to the second when he observes that out-of-equilibrium message. It is easy to rationalize the kind of forward induction used by the receiver in this refinement. Specifically, the greater the sender´s potential profit from deviating to an out-of-equilibrium message is, the stronger his or her temptation to send that message will be. This is the reason why the receiver should attach a higher probability to those types of sender whose potential profit from the deviation is greater. At this point, the reader may be wondering whether those types of sender who might make a greater profit might also be taking a higher risk. However, it will become clear that our new criterion is not attaching a higher probability to those types of sender who assume a higher risk when sending an out-of-equilibrium message in the class of signaling games we analyze in this paper.

This paper is organized as follows. Section 2 will describe the main refinements developed by previous literature. In section 3, some examples will be provided in which those refinements do not limit the number of equilibria and we will introduce our refinement, which will allow us to choose a unique equilibrium. In section 4, we will introduce the general model to which we will apply our refinement, whereas in section 5, we will formally define previous refinements to which we will compare our refinement in monotonic signaling games. In those signaling games, section 6 characterizes the equilibria obtained when the support of the signal is sufficiently narrow and/or the cost of the signal is sufficiently low. After this, section 7 proves that our refinement selects a unique equilibrium when we relax the single-crossing condition, whereas other refinements cannot reduce the number of equilibria. Finally, section 8 summarizes the main conclusions.

2. Literature review

The main concept of equilibrium used to solve signaling games is a perfect Bayesian equilibrium. In particular, a set of strategies chosen by the sender and the receiver and a set of receiver´s beliefs will form an equilibrium whenever they satisfy two conditions. First, the set of strategies must form a Nash equilibrium, that is, each player´s strategy must be a best response to the other player´s strategy. Second, after observing the message chosen by the sender, the receiver must form some beliefs about the type of sender against whom he or she is playing, and those beliefs must be consistent with a Bayesian rule in equilibrium. As the reason why there are multiple equilibria is that the receiver´s beliefs are unrestricted outside the equilibrium, previous literature has proposed different criteria or *refinements* in order to constrain the out-of-equilibrium beliefs. Each of those criteria is a test of the equilibria obtained. If an equilibrium passes the test, we keep that equilibrium, but if the equilibrium fails the test, we discard it.

By and large, each test prescribes a "sensible" way in which the receiver should interpret an out-of-equilibrium message. Imagine that we are testing certain equilibrium. After an out-of-equilibrium message, the receiver should compare the pay-off obtained by each type of sender with that message to the pay-off obtained with an alternative message and using this comparison, the receiver should attach a posterior probability to each type of sender. Each refinement suggests a different comparison or a different way of assigning probabilities to each type of sender using that comparison. Using one of those refinements, an equilibrium will be eliminated whenever the receiver´s beliefs are not consistent with those prescribed by that refinement. In this section, we will briefly describe the main proposals developed by previous literature.

One of the best well-known criteria is the intuitive criterion developed by Cho and Kreps (1987)[1]. We say that an equilibrium will survive the intuitive criterion whenever the out-of-equilibrium beliefs are consistent with the following line of reasoning. If a type of sender can never benefit from a deviation to an out-of-equilibrium message when the receiver responds to this message with any undominated strategies and at the same time, other types of sender

[1] McLennan (1985) suggested using what he called justifiable beliefs, Milgrom and Roberts (1986) used the elimination of the dominated type-message pairs and Kohlberg and Mertens (1986) proposed the "never a weak best response" criterion. Cho and Kreps show that the intuitive refinement gives rise to a lower set of equilibria than those criteria.

can benefit from that deviation, the receiver should attach positive posterior probabilities only to those types of sender who might benefit from the deviation.

Other criteria widely used by previous researchers are the divinity and universal divinity introduced by Banks and Sobel (1987). The divinity criterion will constrain the receiver´s beliefs after a disequilibrium message in the following way. If the set of the receiver´s mixed best responses to that message for which a type of sender, *t*, weakly prefers that message to the equilibrium is a subset of the mixed best responses for which some other type of sender, *t´*, strictly prefers to defect, then the receiver should believe that the sender sending the disequilibrium message is *t´* rather than *t*. Likewise, if the set of the receiver´s mixed best responses to a disequilibrium message, *m*, for which *t* weakly prefers to deviate from a certain equilibrium is a subset of the set of mixed best responses for which any other types of sender strictly prefers to defect, then we would eliminate the *t-m* pair using the universal divinity criterion. It is easy to see that a type-message pair that is eliminated using the intuitive criterion will always be eliminated using the divinity or the universal divinity criterion, but the opposite is not necessarily true. Thus, divinity and universal divinity are stronger refinements than the intuitive criterion[2].

More recently, Fudenberg and He (2018) apply the following line of reasoning to select a particular equilibrium in a signaling game. Let us consider certain equilibrium and an out-of-equilibrium message, *m*. Now, suppose that for every receiver´s strategy such that a *t*-type of sender weakly prefers this message to the best alternative message, another *t´*-type of sender strictly prefers the former message to the latter. In such a case, Fudenberg and He (2018)

[2] In a similar way, Sobel, Stole and Zapater (1990) suggest deleting iteratively the set of weakly dominated strategies, which leads to the concepts of fixed-equilibrium rationalizable outcome (FERO), fixed-equilibrium signal-by-signal rationalizable outcome (FESSO) and fixed-equilibrium rationalizable outcome for the imperfect information game (FERIMO). Sobel et al. (1990) prove that divinity is a stronger refinement than FERO, FESSO y FERIMO because the set of equilibria that survives divinity is a subset of the set of FERO, FESSO and FERIMO.

conclude that the *t´*-type of sender is more compatible with message *m* than the *t*-type. Consequently, after observing message *m*, the ratio of the posterior probability of the *t´*-type to the posterior probability of the *t*-type should be greater than the same prior ratio. This refinement is closely related to the divinity criterion developed by Banks and Sobel (1987), with the only difference that Fudenberg and He (2018) compare each type of sender´s payoff from sending an out-of-equilibrium message to the maximum payoff the sender can obtain with an alternative message, instead of to the payoff received in equilibrium. In fact, in a previous working paper, Fudenberg and He (2017) proved that the set of divine equilibria is a subset of the set of type-compatible equilibria.

A different stream of literature suggests a type of forward induction, which is similar to a fulfilling prophecy. In particular, when a type of sender sends an out-of-equilibrium message, he or she wants to convince the receiver that he or she belongs to a certain set of types. If the receiver chooses his or her optimal response, only those types belonging to that set would prefer that out-of-equilibrium message to the equilibrium. Some criteria eliminate an equilibrium when there exists some set of types who prefer to defect when the receiver interprets an out-of-equilibrium message in this way.

For example, Grossman and Perry (1986) introduced the perfect sequential equilibrium. According to this criterion, if a receiver observes an out-of-equilibrium message, *m*, he should find the set of types of sender who may be better off sending *m* than in equilibrium. If *K* is this set of types of sender, the receiver should concentrate his beliefs on *K* and finally, these should be the only types that are better off sending *m* than in equilibrium. Farrell (1993) developed a similar criterion. He rejects a sequential equilibrium if there exists a set *K* of types who are strictly better off choosing an out-of-equilibrium message when the receiver takes a best response to the belief that he is playing against *K*. An equilibrium for which there is no such out of equilibrium move is called "neologism-proof". More recently, a refinement with the same flavor was suggested by Eso and Schummer (2009). Once again, these authors suggest that the receiver should assume that an out-of-equilibrium message was sent by the only set of types of sender who would benefit from deviating to that message. This should occur whenever the receiver forms any belief over that set of types and take a corresponding best response. Esó and Schummer (2009) consider that the existence of such a message and set of types makes the equilibrium in question not plausible, in which case they say that this equilibrium is vulnerable to a credible deviation.

All previous refinements consider disequilibrium messages as signals, something which was criticized by Mailath, Okuno-Fujiwara and Postlewaite (1993). These authors criticize this type of refinements because they apply only a limited version of forward induction and because the selected equilibrium is discontinuous with respect to the prior distribution of the types of sender in a signaling game. To avoid these problems, they suggest to choose what they call the undefeated equilibria, which are the only equilibria that survive when all players use forward induction to completion. In particular, their test works as follows. Consider an equilibrium and a message, *m*, that is not sent by any player in the equilibrium, but imagine there is an alternative equilibrium in which that message is sent by a non-empty set of types of sender. In such a case, the beliefs at that message in the original equilibrium should be consistent with those beliefs prescribed by the Bayesian rule in the alternative equilibrium. If the beliefs are not consistent with that Bayesian rule, Mailath et al. (1993) consider that the second equilibrium defeats the proposed equilibrium. Therefore, this refinement compares the equilibria obtained in a signaling game and chooses the equilibrium which is undefeated by any other equilibrium.

3. Some examples

In this section, we illustrate the refinement implemented using two simple examples. The first example will be a slightly different version of the beer-quiche signaling game considered by Cho and Kreps (1987), whereas the second example will consider the same game with uncertain pay-offs.

Let us start with the first example illustrated in Figure 1. As in the Cho and Krep´s example, the sender either is wimp or is surly. Nature selects the type of sender and 0.9 is the prior probability that the sender is surly. The prior probabilities are indicated in curly brackets.

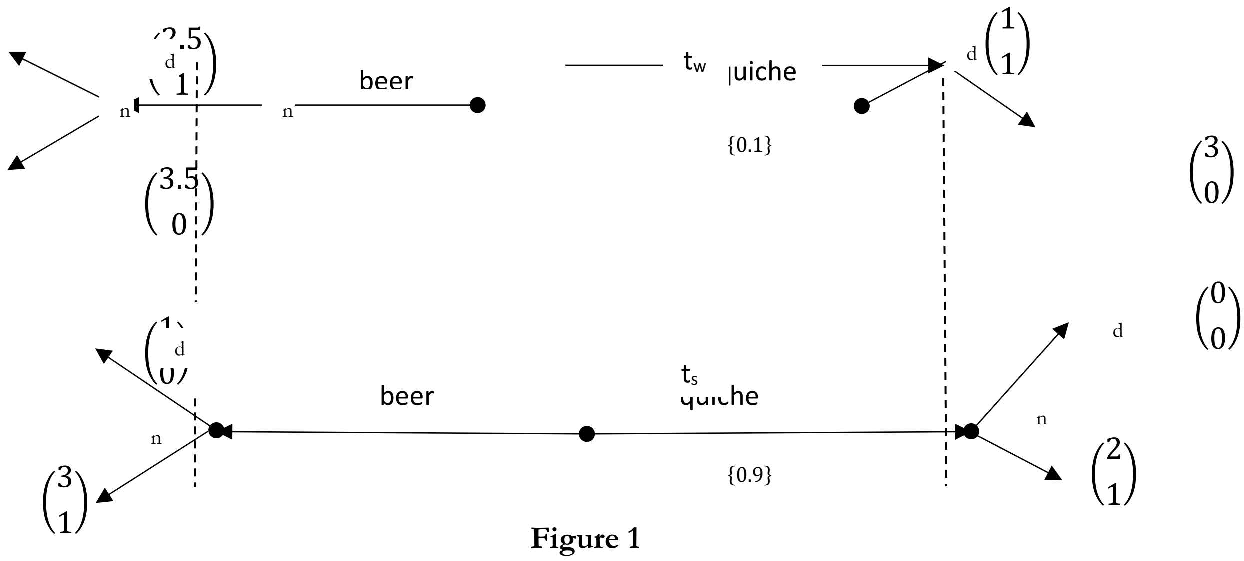


Figure 1

As usual, at the start of the game, the sender knows his type and he has to choose the type of breakfast to have: quiche or beer. Unlike Cho and Kreps´s game, we assume that both types of sender prefer beer for breakfast, but the surly type derives an incremental payoff of 1 from beer, whereas the wimp type´s incremental payoff from having beer is 1.5 if the receiver responds to this breakfast with duel and 0.5 if the receiver responds to it not dueling. After breakfast, the sender meets the receiver, who does not know the sender´s type and has to choose whether to duel (d) or not (n) after observing what the sender had for breakfast. The receiver´s payoffs are the same as those considered by Cho and Kreps (1987) and we also assume that the sender prefers the receiver not to duel. In particular, the sender gets incremental payoff 2 if the receiver chooses not to duel, except when the sender is a wimp and have beer for breakfast, in which case his incremental payoff from not dueling is 1[3]. Finally, the receiver wishes to duel with the sender whenever he is a wimp. The first number in each column vector reflects the sender´s payoff, whereas the second number is the receiver´s payoff.

As in the example introduced by Cho and Kreps (1987), there are two types of pooling equilibria in this game. In the first type of equilibria, both types of sender have quiche for breakfast and the receiver responds to quiche not dueling. To make this a Nash equilibrium, when the receiver observes beer for breakfast, the posterior probability assigned by the receiver to the wimpish type has to be greater than or equal to 0.5. In this type of equilibria, the receiver has to respond to beer with duel or with a behavioral strategy that assigns a probability to dueling greater than 0.5. In the second type of pooling equilibria, both types of sender have beer for breakfast and the receiver responds to beer by not dueling. This Nash equilibrium is independent of the posterior probability attached by the receiver to the wimpish type when quiche is observed. In this type of equilibria, the receiver might respond to quiche attaching any probability, μ_2, to dueling whenever $0 \leq \mu_2 \leq 1$.[4]

Now, we will demonstrate that both types of equilibria survive the refinements introduced by previous literature and a different criterion will be introduced in order to select only one of those equilibria.

---

[3] To rationalize the numbers, we might imagine that the wimpish type gets drunk easily when he drinks beer and a duel will be less painful for him when he is drunk than when he is not (having quiche for breakfast).

[4] There is also another type of equilibria in mixed strategies, in which the wimpish type of sender plays $(p, 1-p)$, while the surly type plays $(q, 1-q)$, the receiver responds $\left(\frac{1}{2}, \frac{1}{2}\right)$ to beer and not dueling to quiche, where p is the probability assigned by the wimpish type to beer in this type of equilibria, $1-p$ is the probability assigned by the wimpish type to quiche, q is the probability assigned by the surly type to beer and $1-q$ is the probability assigned to quiche. Notice that the receiver responds to beer assigning a probability of 0.5 to dueling and 0.5 to not dueling. Additionally, if the receiver observes beer for breakfast, the posterior probability of a wimpish type will be equal to 0.5 and finally, if the receiver observes quiche for breakfast, the posterior belief of a wimpish type has to be lower than 0.5 so that the receiver has incentives to respond to quiche by not dueling. However, we will only focus on the pure strategy equilibria for the sake of simplicity.

Intuitive criterion. It is easy to see that both types of equilibria pass the intuitive criterion. In particular, in the first type of equilibria, where both types of sender choose quiche for breakfast and the receiver responds to it by not dueling, the out of equilibrium message, which is beer, is not dominated in equilibrium for both types of sender. For example, the wimp will be better off choosing beer than in equilibrium whenever the receiver responds to beer not dueling. The same occurs for the surly. In the second type of equilibria, where both types of sender choose beer and the receiver responds to it by not dueling, the out of equilibrium message, which is quiche, is dominated in equilibrium for both types of sender. In fact, both types of sender will get a lower pay-off choosing quiche than in equilibrium whatever the receiver´s response to quiche. Therefore, both types of equilibria survive the intuitive criterion.

Divinity and universal divinity. In the first type of pooling equilibria with quiche for breakfast, the wimp will prefer to deviate to beer whenever the receiver responds to beer with a probability of dueling lower than 0.5. Then, the wimpish type is indifferent between the pay-off obtained in equilibrium and deviating to beer when the probability assigned by the receiver to dueling in response to beer is equal to 0.5. The same will occur for the surly. Thus, as the set of the receiver´s mixed best responses that makes each type of sender wish to defect and send beer or be indifferent between quiche and beer are the same, the equilibrium survives the divinity criterion. As there are only two types of sender in this example, the universal divinity criterion coincides with divinity.

In the second type of pooling equilibria with beer for breakfast, none of the types of sender will be able to be better off by deviating to quiche whatever the receiver´s response to that deviation. Thus, this equilibrium also survives the divinity and universal divinity refinements.

Obviously, as the divinity criterion is stronger than the criteria introduced by McLennan (1985), Milgrom and Roberts (1986), Kohlberg and Mertens (1986), Sobel, Stole and Zapater

(1990) and Fudenberg and He (2018), both types of equilibria considered also survive all of these refinements.

Perfect sequential equilibria and similar refinements. It is straightforward to show that both types of sequential equilibria in this game survive the refinements suggested by Grossman and Perry (1986) and Farrell (1993). Let´s start with the equilibrium with quiche for breakfast. In this equilibrium, if the receiver observes beer, it is "credible" to assume that it was the surly type who sent that message. In fact, if the surly type sends this message, the receiver prefers not to duel and the surly type will get 3 rather than 2. However, if the receiver responded to beer not dueling, the wimp type would also prefer beer rather than quiche. Thus, it is not possible to discard this equilibrium using the restriction imposed by this notion of equilibrium. Similarly, in the second type of equilibria where both types of sender choose beer for breakfast, no type can be better-off sending the out-of-equilibrium message and for this reason, the posterior beliefs after quiche are unrestricted. For the same reasons, both equilibria are not vulnerable to a credible deviation as suggested by Esó and Schummer (2009)[5].

Undefeated equilibrium. If we only considered the equilibria in pure strategies of the signaling game in Figure 1 previously considered, we could argue that the only equilibrium surviving the criterion developed by Mailath et al. (1993) in this example is the pooling equilibrium in which both types choose beer for breakfast. In fact, the pooling equilibrium with quiche would be defeated by the equilibrium with beer for breakfast. According to this refinement, in the quiche-equilibrium, the receiver should attach a probability of 0.9 to the surly type after beer (the out-of-equilibrium message) because this is the probability attached to the

[5] In the example introduced by Esó and Schummer (2009) in their Figure 1, the only equilibrium that survives their criterion is also the only equilibrium that survives our criterion of the most profitable deviator. Kohlberg and Mertens (1986) use a similar signaling game in their Figure 14 in which there are two types of equilibria that previous solution concepts cannot distinguish, i.e., they are perfect, proper, proper after elimination of dominated strategies, etc. However, there is only one type of equilibria which is stable and survives the intuitive criterion and this is also the only type of equilibria which survives our most profitable deviator criterion.

surly type after beer in the alternative equilibrium. Therefore, given this belief, the receiver should choose not to duel after beer, which will make both types of sender defect from quiche to beer. However, there is an alternative equilibrium in mixed strategies as shown in footnote 2. As the receiver responds to beer by dueling with a probability of 0.5 and by not dueling with a probability of 0.5, the posterior beliefs associated to this response are also possible after beer, in which case both types of sender make no profit from deviating to beer and consequently, the pooling equilibrium with quiche for breakfast is actually not defeated. As shown by Mailath et al. (1993), a unique posterior probability can be attached to a particular unsent message at a particular equilibrium whenever there is only one alternative equilibrium in which this out-of-equilibrium message was sent. If there are several alternative equilibria in which a given disequilibrium message is sent, the receiver will not be able to make an unambiguous comparison between the proposed equilibrium and possible alternatives. For this reason, Mailath et al. (1993) propose to require that there be a unique alternative equilibrium in which the given disequilibrium message is played and which has the additional properties in the definition of undefeated equilibrium[6].

New criterion. Now, a new refinement will be developed, which will allow us to choose only one type of equilibria in this type of models. In this refinement, we will test an equilibrium as follows. We assume that a type of sender would wish to send an out-of-equilibrium message whenever he or she can make a profit from that deviation. Furthermore, the higher the potential profit from the deviation is, the stronger the sender´s temptation to deviate from the equilibrium. Consequently, if the receiver observes an out-of-equilibrium message, he or she will assign the type of sender who might make the highest possible profit from that

[6] Mailath et al. (1993) also provide an example in which there are four equilibria in pure strategies, but there is no undefeated equilibria. In particular, each of those equilibria is defeated by another one. However, only one of those equilibria survives our most profitable deviator criterion, which selects a unique equilibrium in that game.

deviation a higher posterior probability than the prior probability[7]. Let us use the first type of pooling equilibria to illustrate this refinement, which will be called the *most profitable deviator*. In this equilibrium, both types of sender choose quiche and the receiver does not duel after observing this message. Thus, the wimp is getting 3 in equilibrium, whereas the surly is getting 2. If the wimp chooses beer for breakfast (an out-of-equilibrium message), the best he can get from this deviation is 3.5 provided that the receiver responds by not dueling. It means that the wimp´s highest possible profit from this deviation will be 0.5. Similarly, if the surly chooses beer for breakfast, the best he can get from this deviation is 3 giving rise to a profit of 1. As the surly can make a greater profit from sending beer than the wimp, the receiver should assign a greater probability to the surly type than the prior probability (0.9) after beer. In such a case, the receiver would respond to beer by not dueling and both types of sender would prefer beer to quiche. In conclusion, the pooling equilibrium in which both types of sender choose quiche does not survive our criterion of the most profitable defector.

At this point, some readers may be skeptical about this line of reasoning arguing that the surly may also be the type of sender who is taking a higher risk when deviating to beer. As we are assigning a higher posterior probability to the type of sender who can make a greater profit from a deviation regardless of the risk assumed, some readers might claim that we are assuming that the sender is a risk lover. In order to calm those readers who might be worried about this strong assumption, we will show them that our assumptions about the sender´s risk preferences are sensible in the type of signaling games we are analyzing here. Before undertaking this task, it is necessary to remind the skeptical reader that our main purpose is not to provide a superior criterion that can be applied in any signaling games, but to provide a new criterion which is useful when previous refinements do not work. Therefore, now, we will compare our criterion to other criteria to show the reader that in the class of signaling

[7] Fudenberg and Tirole (1991) suggests that there is no intuitive justification for the receiver to put infinitely more weight on sender types that gain from a deviation more often than others. We follow this line of reasoning.

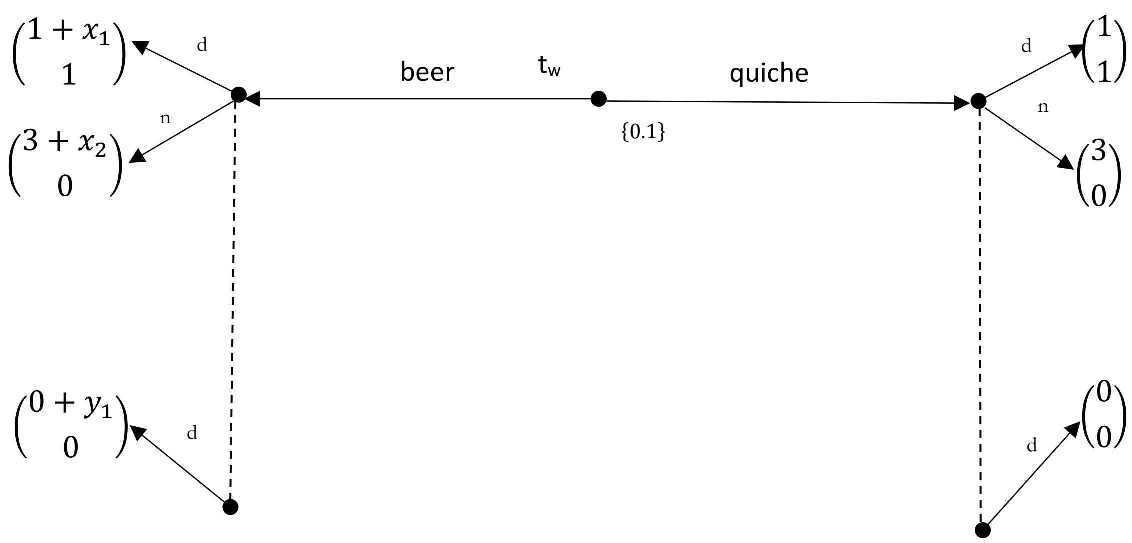
games we are analyzing, we are not making any insensible assumption about the sender´s risk preferences.

First, let us compare our criterion to the divinity criterion in those games where the latter does not work, which is the context in which our criterion should be used. Given an equilibrium, imagine that I_t^m denotes the set of the receiver´s best responses to a disequilibrium message, *m*, that makes a *t-type* of sender indifferent or better off than in equilibrium. Similarly, $B_{t'}^m$ may represent the set of the receiver´s best responses to *m* that makes the *t´-type* of sender strictly better off than in equilibrium. According to the divinity criterion, if $I_t^m \subset B_{t'}^m$, the receiver should discard the *t-type* of sender after observing *m*. This is equivalent to discarding the *t-type* of sender after observing *m* whenever $\overline{B_{t'}^m} \subset \overline{I_t^m}$, where $\bar{S}$ means the complement of the set S. Thus, the receiver should discard the *t-type* after *m* whenever those mixed best responses to *m* which make the *t´-type* indifferent or worse off than in equilibrium are a subset of those mixed best responses which make the *t-type* strictly worse off than in equilibrium. In other words, the receiver discards the *t-type* after *m* because this type of sender is assuming a higher risk with such a deviation than the *t´-type*. Here, we measure the risk assumed by a *t-type* of sender when sending a disequilibrium message, *m*, as the set of the receiver´s best responses to *m* that makes the *t-type* worse off than in equilibrium. As the set of those mixed best responses to *m* for which the *t-type* strictly loses money when sending that disequilibrium message, which is represented by $\overline{I_t^m}$, is wider than the set of those mixed best responses to *m* for which the *t´-type* weakly loses money, represented by $\overline{B_{t'}^m}$, the risk assumed by the *t-type* of sender when deviating to *m* is higher than the risk assumed by the *t´-type*. This is the reason why the divinity criterion would discard the *t-type* after observing *m*. However, we cannot rule out the pooling equilibrium with quiche using the divinity criterion because we cannot identify a type who is assuming more risk than the other when sending beer. In this context, if no type assumes more risk than another when deviating to beer, we argue that the receiver should assign a greater posterior probability to the type of sender who can make a greater profit.

There might also be some readers who prefer to interpret an out-of-equilibrium message as a message sent by some type of sender in an alternative equilibrium. Let us compare our criterion to the undefeated criterion introduced by Mailath et al. (1993) in order to show the relationship between our criterion and other criteria with that flavor. When we applied the undefeated criterion to test the pooling equilibrium with quiche for breakfast, we argued that there are two alternative equilibria in which some type of sender sends the out-of-equilibrium message, beer. As the receiver´s posterior beliefs after observing beer in the quiche-equilibrium are consistent with the posterior beliefs in one of those alternative equilibria, we could not rule out the quiche-equilibrium. In one of the alternative equilibria, which is an equilibrium in mixed strategies, the receiver assigns the same probability to dueling as to not dueling after beer, whereas in the other the receiver responds to beer by not dueling. We know that the wimpish type´s pay off in the quiche equilibrium is 3, whereas this type of sender will also receive 3 if he sends beer in the equilibrium in mixed strategies or 3.5 in our alternative pooling equilibrium in pure strategies. Likewise, the surly type´s pay off in the quiche equilibrium is 2, whereas this type will also receive 2 if he sends beer in the equilibrium in mixed strategies, or 3 in the equilibrium in pure strategies. Therefore, assuming that beer must be interpreted as a signal of an alternative equilibrium, none of the types of sender would be taking risk from such a deviation. In this context, it seems sensible to propose our criterion: The receiver should believe that the surly type is more likely to deviate from the quiche equilibrium by sending beer. In fact, this type of sender can make a greater profit from that deviation to an alternative equilibrium on average, but this type assumes the same risk than the other.

However, it is easy to see that our second pooling equilibrium, in which both types of sender choose beer and the receiver responds to this breakfast by not dueling, survives our criterion of the most profitable deviator. In fact, none of the types of sender can be better off by sending the out-off-equilibrium message whatever the receiver´s response to the deviation.

To finish this section, we will consider another example with uncertain pay-offs in order to prove the power of our refinement in a wide variety of situations. Figure 2 illustrates this new signaling game, which is similar to the game considered in Figure 1, but there are only two differences. First, the wimp type always prefers beer to quiche, but the additional pay-off from choosing beer is unknown by the receiver. The only information that the receiver has is that this additional pay-off is a random variable, X, whose marginal distribution function is $F(X)$, which is common knowledge. In Figure 2, x_1 and x_2 are two realizations of this distribution function. Second, the surly type also prefers beer to quiche, but once again, the additional pay-off from choosing this type of breakfast is unknown by the receiver. Now, this additional pay-off is a random variable, Y, whose marginal distribution function is $G(Y)$, and the joint distribution function of both variables is $H(X,Y)$. Both distributions are common knowledge. In Figure 2, y_1 and y_2 are only two realizations of the distribution of Y. We also assume that $0 < X, Y < 2$.



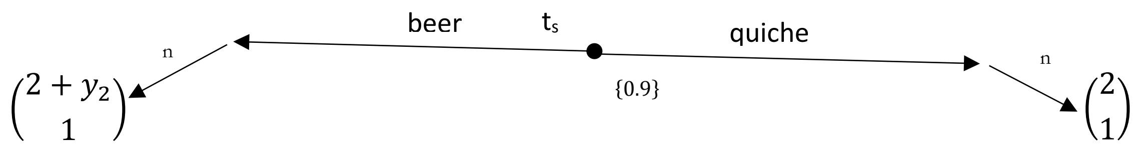


Figure 2

In this game with random payoffs, there are also two types of pooling equilibria. Once again, in the first type of equilibria, both types of sender will choose quiche for breakfast, whereas the receiver will respond to beer by dueling and he will respond to quiche by not dueling. In this equilibrium, the receiver´s posterior probability of facing a wimpish type after observing beer has to be greater than 0.5, so that it will be rational to choose to duel after beer. However, the posterior probability assigned to a wimpish type after observing quiche will be equal to the prior, which is 0.1, as prescribed by the Bayesian rule.[8]

In the second type of pooling equilibria of this game, both types of sender will choose beer for breakfast, after which the receiver will choose not to duel. This type of equilibria will be consistent with any receiver´s answer after quiche and any posterior belief after quiche. As both types of sender choose beer for breakfast, the receiver´s posterior beliefs of each type of sender after observing beer will be equal to the prior[9].

Using the same arguments as in the game of Figure 1, we conclude that both types of equilibria considered in this alternative game also pass the intuitive criterion and are perfect sequential equilibria as defined by Grossman and Perry (1986) and "neologism-proof" as defined by Farrell (1993). In this game, it is not possible to see whether our first type of pooling equilibria passes the divinity criterion or not. To do so we need to identify the

[8] There will be another pooling equilibrium in which both types of sender choose quiche for breakfast, the receiver responds to quiche by not dueling, but he responds to beer with a mixed behavioral strategy. In this equilibrium, the posterior probability assigned by the receiver to the wimpish type after beer will be 0.5, but the posterior probability assigned to the wimp after quiche will be 0.1.

[9] Ex-post, that is, after the realizations of *X* and *Y*, another pooling equilibrium in "quasi" mixed strategies would arise. In that equilibrium, the wimp will choose beer, but the surly will choose beer with probability equal to 0.1 and quiche with probability 0.9. Finally, the receiver will respond to quiche by not dueling, while he will respond to beer by dueling with probability μ_1 and choosing not to duel with probability $1-\mu_1$, where $\mu_1 = \frac{y_2}{2+y_2-y_1}$. The receiver´s posterior belief of a wimp after beer will be equal to 0.5, but the same posterior belief after quiche will be equal to 0.

receiver´s responses that makes each type of sender better off when they choose the out-of-equilibrium message, which is beer. As the sender´s payoffs from sending beer are not perfectly observed by the receiver, it does not seem sensible to use the receiver´s responses to beer that make each type of sender better off to build the receiver´s beliefs after beer. In spite of these difficulties, the second type of equilibria considered also pass the divinity criterion for the same reasons as those used in Figure 1.

At this point, it will be necessary to see how we can use our criterion to rule out the first type of pooling equilibria. Imagine that the wimp deviates from the quiche-equilibrium by sending beer. In this case, his maximum profit from this deviation will be x_2. However, if the surly deviates from quiche in our equilibrium to beer, his maximum possible profit from doing so is y_2. As shown before, the receiver does not observe those realizations of the random variables, X and Y, and it will be strange to base his beliefs on those realizations. In spite of this problem, it is not difficult to consider situations in which the correlation between the types of sender and the random payoffs is known[10]. For example, imagine that the surly is more likely to obtain a greater payoff from beer than the wimp. In other words, imagine that the distribution of Y first-order stochastically dominates the distribution of X. To simplify the explanation, we can assign a numerical value to each type of sender through the following function $T: \{t_w, t_s\} \longrightarrow [0,1]$. For example, imagine that $T(t_w) = 0, T(t_s) = 1$. Using this notation, another way of expressing our additional assumption related to stochastic dominance is saying that the random payoffs are positively correlated with the type of sender. In such a case, the surly type is more likely to have a greater profit from beer than the wimp. Under this assumption, we can use the most profitable deviator criterion in order to rule out

[10] For example, in the signaling game of the labor market introduced by Spence (1974), we may assume that the cost of education incurred by each worker is unknown by the employer, but it is negatively correlated with the sender´s type. It would mean that the cost of education, which might include monetary and psychological costs, tend to decrease with each worker´s ability. In particular, one might assume that each worker might be subject to negative random shocks (possibly psychological shocks), which might increase their cost of education, but those workers with a greater ability will have more strategies at their disposal to adapt to those negative shocks.

the first type of pooling equilibrium. Specifically, when the surly type deviates from this equilibrium choosing beer for breakfast, he will be more likely to make a "high" profit than the wimp. If the correlation between the types of sender and the random payoffs previously described is common knowledge, the posterior probability assigned by the receiver to the surly after observing beer must be greater than the prior. In such a case, the receiver will not duel after beer and consequently, both types will prefer beer to quiche. In conclusion, our first equilibrium does not survive our most profitable deviator refinement. It is easy to see that this equilibrium would survive this refinement if the correlation between the random payoffs and the types of sender were negative rather than positive. Obviously, the second type of pooling equilibria with beer for breakfast survives the most profitable deviator criterion whatever the type of correlation considered.

To sum up, we have introduced a different refinement to choose a pooling equilibrium in a general class of signaling games in which previous criteria do not work. In figures 1 and 2, we illustrate two examples of those games. Let us use our pooling equilibria with quiche for breakfast in both games in order to identify the main characteristics of the class of games we are analyzing. For example, we can observe that the wimp will be worse off by deviating to beer if the receiver discovers that he is playing against a wimp after beer. In this respect, we can deduce that beer is a costly signal. However, if sending beer makes the receiver believe that the sender belongs to a higher set of types, that is, if the receiver believes that the sender is a surly, the wimp will make a profit from such a deviation. In this respect, the cost of this signal (sending beer) is sufficiently low. Hence, in this class of signaling games in which either the support of the signal is sufficiently narrow or the cost of the signal does not decrease too much with the type of sender, there will be multiple pooling equilibria and none of the refinements introduced by the previous literature can select a unique equilibrium. We found that our criterion, the most profitable deviator, is useful to choose a particular equilibrium in this general class of signaling games.

Our refinement is stronger than those previously suggested. In fact, there might be signaling games in which no equilibria survives our criterion. However, the main problem of most signaling games studied is the multiplicity of equilibria rather than the inexistence of one equilibrium.

4. The model

Now, we can introduce the main elements of our model. In this paper we will consider a generic class of signaling games with finite strategy sets. There are two players, a Sender (*S*) and a Receiver (*R*). The Sender has private information, summarized by his type, *t*, an element of a finite set, *T*. Specifically, $T = \{t_1, t_2, \cdots, t_n\} \subset \mathbb{R}$. For notational convenience, it is assumed that $t_1 < t_2 < \cdots < t_n$. There is a strictly positive probability distribution *p(t)* on T, where *p(t)* is the ex ante probability that *S*´s type is *t* and is common knowledge. After observing his type, *S* sends a message, *m*, to R, where *m* is an element of a finite set, $M(t) = \{m_1, m_2, \cdots m_N\}$. After observing *m*, R chooses an action, *a*, from a finite set *A(m)*. The sender and the receiver have von Neumann-Morgenstern utility functions *u(t, m, a)* and *v(t, m, a)*, respectively.

In this model, the strategy chosen by the sender will be a function $q: T \longrightarrow \Delta_{\bar{M}}$, where $q(m|t)$ is the probability that *S* sends the message *m*, given that his type is *t*, and $\Delta_{\bar{M}}$ denotes the set of possible mixed strategies that the sender can choose. Likewise, the strategy chosen by the receiver will be a function $r: M(t) \longrightarrow \Delta_{\bar{A}}$, where $r(a|m)$ is the probability that R chooses the action *a*, given that he has observed the message *m*, and $\Delta_{\bar{A}}$ denotes the set of possible mixed strategies that the receiver can choose. After observing the message *m*, the receiver will have to form posterior beliefs about the types of sender who sent that message. In particular, the receiver´s posterior beliefs upon receiving the sender´s message is a function $\mu: M(t) \longrightarrow \Delta_{\bar{T}}$, where $\mu(t|m)$ is the posterior probability assigned by *R* to *t* after observing *m* and $\Delta_{\bar{T}}$ is the set of probability distributions on *T*.

As usual, we restrict our attention to perfect Bayesian equilibria, which are the set of signaling rules for *S*, $q(m|t)$, action rules for *R*, $r(a|m)$, and beliefs for *R*, $\mu(t|m)$, that satisfy the following conditions:

1. Rationality: The strategy chosen by one player will maximize its utility given the strategy chosen by the other player:

    i. $\forall t \in T, q(m^*|t) > 0$ only if $\bar{u}[t, m^*, r(a|m^*)] = \max_{m \in M(t)} \bar{u}[t, m, r(a|m)]$.

    ii. $\forall m \in M(t), r(a^*|m) > 0$ only if $\sum_{t \in T} v(t, m, a^*)\mu(t|m) = \max_{a \in A(m)} \sum_{t \in T} v(t, m, a)\mu(t|m)$.

    Where $\bar{u}[t, m, r(a|m)] = \sum_{a \in A(m)} u(t, m, a) r(a|m)$.

2. Consistency: The receiver´s posterior beliefs have to be consistent with the Bayesian rule in equilibrium: If $\sum_{t \in T} q(m|t)p(t) > 0$, then $\mu(t^*|m) = \frac{q(m|t^*)p(t^*)}{\sum_{t \in T} q(m|t)p(t)}$.

5. Equivalence of refinements in monotonic signaling games

In this section, we introduce previous refinements in a formal way and introduce our first assumption, under which all previous refinements are equivalent.

Now, in order to define some criteria previously used and our particular refinement, let $MBR(\mu, m)$ be the set of mixed strategies that are best responses to *m* given the assessment μ. Likewise, let $MBR(I, m)$ represent the corresponding set of mixed strategies that are best responses by the receiver to assessments concentrated on the subset *I* of *T*, whereas $MBR(t, m)$ represents the mixed best response by the receiver to *m* given a posterior belief concentrated on the *t*-type of sender. Finally, we will write $MBR_p(T, m)$ to denote the set of mixed best responses to message *m* in a pooling equilibrium in which all types of sender are choosing that message and the receiver´s posterior beliefs are equal to the prior probabilities over types. We assume that all those mixed best responses exist for each value of *m* and for each set of posterior beliefs because *v(t, m, a)* is quasiconcave with respect to *a*. Then, fix a sequential equilibrium (q, r, μ) and let $\bar{u}^*(t)$ be the equilibrium expected utility of the *t-type* of sender. Choosing an off-the-equilibrium-path signal *m*, we can define:

$$P(t|m) = \{r \in MBR(T, m): \bar{u}^* < \bar{u}(t, m, r)\} \quad (1)$$

And

$$P^0(t|m) = \{r \in MBR(T, m): \bar{u}^* = \bar{u}(t, m, r)\} \quad (2)$$

As usual, $P(t|m)$ is the set of best responses of the receiver to *m* that gives incentives to the *t-type* of sender to deviate from his equilibrium strategy, whereas $P^0(t|m)$ is the set of best responses of the receiver to *m* that gives the *t-type* of sender the same utility as in equilibrium.

To start with, we can define the criterion, called divinity or D1 in Cho and Kreps (1987) and in Cho and Sobel (1990). This refinement rules out strategy *m* of the *t-type* of sender if there exists $t´$ such that

$$P(t|m) \cup P^0(t|m) \subset P(t´|m) \quad (3)$$

A sequential equilibrium survives criterion D1 if and only if, for all off-the-equilibrium-path messages *m*, $\mu(t|m) = 0$ whenever (3) holds for some $t´$ such that $P(t´|m) \neq \emptyset$. Additionally, types *t* for which (3) does not hold are also said to survive criterion D1.

Similarly, Banks and Sobel (1987) show that universal divinity eliminates strategy *m* of the *t-type* of sender if:

$$P(t|m) \cup P^0(t|m) \subset \cup_{t \neq t´} P(t´|m) \quad (4)$$

The criterion of never a weak best response developed by Kohlberg and Mertens (1986) and Cho and Kreps (1987) is stronger than universal divinity. This test eliminates strategy *m* of the *t-type* of sender if:

$$P^0(t|m) \subset \cup_{t \neq t'} P(t'|m) \quad (5)$$

To compare our refinement to previous criteria, we will restrict our attention to monotonic signaling games. As shown by Cho and Sobel (1990) a game is monotonic, whenever the following condition is satisfied:

C1. For all $m \in M$ and $r, r' \in MBR(T, m)$, if $\bar{u}(s, m, r) > \bar{u}(s, m, r')$ for some $s \in T$, then $\bar{u}(t, m, r) > \bar{u}(t, m, r')$ for all $t \in T$.

The comparison between our refinements and the previous refinements will be easier in this type of games because criterion D1, universal divinity and never a weak best response are equivalent in monotonic signaling games (Cho and Sobel, 1990). Furthermore, there might be some non-monotonic signaling games in which the refinement introduced by Grossman and Perry (1986) is stronger than ours[11], but we will only consider those signaling games which satisfies C1. In fact, in monotonic signaling games, the perfectness introduced by Grossman and Perry (1986) does not restrict the number of equilibria. In this class of games, if a type of sender can be better-off sending an out-of-equilibrium message given a particular best response chosen by the receiver, the other types of sender will also prefer this type of best response to the same out-of-equilibrium message. Therefore, given the out-of-equilibrium message, any beliefs about the types of sender who sent that message will be "credible" according to the definition of the perfect sequential equilibrium. For similar reasons, a sequential equilibrium in a monotonic signaling game will always pass the criterion suggested by Farrell (1993).

[11] For example, see the games of figures 2 and 4 in Grossman and Perry (1986).

6. Potential equilibria with a limited set of signals

In this section we show the relationship between our notion of equilibrium and those refinements developed by the previous literature in a general class of monotonic signaling games in which the usual single-crossing condition is satisfied and the set of signals is limited. Before presenting our first result, we will formalize our refinement using the following definition.

Definition 1 (Most profitable deviator). Given an equilibrium, $\left(q(m|t), r(a|m), \mu(t|m)\right)$, when the receiver observes an out-of-equilibrium message, *m*´, the posterior probability must satisfy the following condition:

$$\frac{\mu(s|m')}{\mu(s'|m')} > \frac{p(s)}{p(s')} \quad \text{whenever} \quad \max_{r\in MBR(T,m')} \bar{u}[s, m', r(a|m')] - \bar{u}^*(s) > \max_{r\in MBR(T,m')} \bar{u}[s', m', r(a|m')] - \bar{u}^*(s) \quad \forall s' \neq s \tag{6}$$

After observing an out-of-equilibrium message, this condition implies that the receiver should assign a posterior probability greater than that prescribed by the prior beliefs to the type of sender who can obtain the highest possible benefit from that deviation. In this definition we do not impose a unique value for the receiver´s posterior beliefs after observing the out-of-equilibrium message. Doing so would imply some kind of forward induction that must be common knowledge.

There might be some situations in which an equilibrium passing our refinement fails to exist[12]. In such a case, a weaker refinement will suffice.

Now, we will introduce some notation:

[12] For example, in the game suggested by Grossman and Perry (1986) in their Figure 7, no equilibria survive our refinement.

Definition 2 (Strict order of two sets of types): Let $T_1, T_2 \subseteq T$ be two sets of types of sender, such that $T_1 = \{t_1^1, t_1^1, \cdots, t_k^1\}$ and $T_2 = \{t_1^2, t_2^2, \cdots, t_p^2\}$. We will say that $T_1 < T_2$ provided that, at least, one of the following conditions is satisfied:

i. If $min(T_1) = min\{t \in T_1\} < min(T_2) = min\{t \in T_2\}$ and there exists a type, $t_j^2 \in T_2$, such that $t_j^2 = t_i^1$ whenever $t_i^1 \in T_1$ and $t_i^1 > min(T_2)$.

ii. If $max(T_2) = max\{t \in T_2\} > max(T_1) = max\{t \in T_1\}$ and there exists a type, $t_i^1 \in T_1$, such that $t_i^1 = t_j^2$ whenever $t_j^2 \in T_2$ and $t_j^2 < max(T_1)$.

Roughly speaking, we say that a set of types is higher than another set when the former includes higher types than the latter.

Now, we specify the class of signaling games in which our refinement will be more binding than previous refinements. In particular, we consider the following assumptions:

C2. $\bar{u}(t, m, MBR(T_1, m)) < \bar{u}(t, m_N, MBR(T_2, m_N))\ \forall t \in T,\ \forall m \in \{m_1, m_2, \cdots m_N\}$ and $\forall T_1, T_2 \subseteq T$, such that $T_1 < T_2$.

C3. $\bar{u}(t, m_N, MBR(T', m_N)) < \bar{u}(t, m, MBR(T', m)\ \forall t \in T, \forall T' \subseteq T, \forall m < m_N$.

Where $\bar{u}(t, m, MBR(T_1, m))$ is the level of utility that the *t*-sender will obtain when he chooses *m* and the receiver chooses its mixed best response to *m* given the fact that the posterior probabilities are concentrated on the types belonging to T_1.

Assumption C2 means that any type of sender would prefer to make the receiver believe that he belongs to a higher set of types even if this type of sender had to send the highest possible message to do so. In other words, the support of the signal is sufficiently narrow and/or the cost of the signal is sufficiently low.

Assumption C3 means that any type of sender will prefer to send a lower message than the highest possible message whenever the receiver does not change his posterior beliefs. Thus, this assumption implies that the cost of the signal is not too low.

Now, we can characterize the type of equilibria obtained under assumptions C2 and C3. Under these assumptions, the next lemma proves that there will be infinite pooling equilibria. In each pooling equilibrium, all types of sender will send the same message, which can be any of the messages belonging to $M(t)$.

LEMMA 1. *In a signaling game in which C2 and C3 are satisfied, there will be multiple pooling equilibria. Each pooling equilibrium will be defined by a particular message sent by all types of sender and that message can take any message of* $M(t) = \{m_1, m_2, \cdots m_N\}$.

Proof. Let us consider one of those potential equilibria in which all types are sending the same message, which can be any $m \in \{m_1, m_2, \cdots m_N\}$, the receiver responds to that message using his mixed best response, $MBR_p(T, m)$, given the fact that the prior and posterior probabilities are the same. We will prove that any of those outcomes will be a sequential equilibrium.

If we apply C3 when $T' = T$, the following inequality is obtained:

$$\bar{u}\big(t, m, MBR_p(T, m)\big) \geq \bar{u}\big(t, m_N, MBR_p(T, m_N)\big) \; \forall t \in T, \forall m \leq m_N \quad (7)$$

Obviously, if $m = m_N$, the previous inequality will become an equality. Now, using C2, we obtain a lower bound for the right-hand side of inequality (7):

$$\bar{u}\big(t, m_N, MBR_p(T, m_N)\big) > \bar{u}\big(t, m', MBR(T_1, m')\big) \; \forall t \in T, \forall m' \neq m_N, \forall T_1 < T \quad (8)$$

In (8) T_1 may be any set of types lower than T. For example, it may be a set consisting of the first *p* types of sender, where $p < n$. Then, if the receiver´s posterior beliefs are sufficiently concentrated on lower types after observing an out-of-equilibrium message, any type of sender would not want to deviate from a pooling equilibrium sending a different message. As the out-of-equilibrium beliefs are unrestricted in a perfect Bayesian equilibrium, there will be out-of-equilibrium beliefs that will sustain any pooling equilibrium. Consequently, each type of sender´s choice will be individually rational. Additionally, by construction, the receiver´s choice is individually rational because it is a best response to the sender´s message when the posterior and the prior beliefs are equal. Finally, as it is a pooling equilibrium, according to the Bayesian rule, the posterior beliefs have to be equal to the prior beliefs. Therefore, those beliefs are consistent. We have proven that there will be N pooling equilibria in our class of signaling games satisfying C2 and C3. QED.

Our next lemma discards some potential equilibria in the class of games we are analyzing.

LEMMA 2. *In a signaling game in which C2 and C3 are satisfied, there cannot be an equilibrium in pure strategies in which all the types belonging to a set of types, $T_l \subset T$, are sending the same message, but this message is different from that sent by all types of sender belonging to a different set, $T_h \subset T$, if $T_l < T_h$.*

Proof. Using this lemma, we discard all completely separating equilibria[13] and many semipooling equilibria. In order to prove this lemma, imagine there were an equilibrium in pure strategies in which all types belonging to T_l are sending *m*, whereas those types belonging to T_h are sending *m´*.

In this case, $T_l \cap T_h = \emptyset$ because it is an equilibrium in pure strategies. Let us call $t_{max} = max(T_l)$. We can apply C2 to $T_1 = T_l$ and $T_2 = T_h$ in order to obtain:

$$\bar{u}(t_{lmax}, m, MBR(T_l, m)) < \bar{u}(t_{lmax}, m_N, MBR(T_h, m_N)) \quad (9)$$

C2 can be applied because $T_l < T_h$. Now, we apply C3 and obtain:

$$\bar{u}(t_{lmax}, m_N, MBR(T_h, m_N)) \leq \bar{u}\left(t_{lmax}, m', MBR(T_h, m')\right) \quad (10)$$

[13] A completely separating equilibrium is an equilibrium in which each type of sender sends a different message.

Obviously, if $m' = m_N$, inequality (10) will become an equality. Thus, we have proven that there would be a type of sender, t_{lmax}, who would prefer to choose the message of the other set of types. In the separating or semipooling equilibrium we are testing, the receiver´s posterior beliefs should be consistent with the Bayesian rule. For this reason, after observing m', the receiver should believe that only those types of sender belonging to T_h could have sent that message. According to our assumptions, those beliefs will make a type of sender who belongs to T_l wants to deviate. Consequently, a group of types cannot send a different message from that sent by a group of types who are strictly lower in a sequential equilibrium if assumptions C2 and C3 are satisfied. QED.

Lemmas 1 and 2 lead to the next result.

THEOREM 1. *In a signaling game in which C2 and C3 are satisfied, there will be multiple sequential equilibria in pure strategies. Those equilibria can only be pooling and semipooling equilibria. In the latter case, there must be at least a type of sender sending a different message between a pair of types sending the same message.*

Proof. The proof of this theorem is straightforward from Lemmas 1 and 2.

Interestingly, we did not need to use assumption C1 in order to prove Lemmas 1 and 2. Therefore, Theorem 1 characterizes the type of equilibria obtained when assumptions C2 and C3 are satisfied irrespective of whether the game is monotonic or not.

7. Equilibrium selection

In this section, we will analyze the type of single-crossing condition that will allow us to choose a unique equilibrium in the class of signaling games in which C1, C2 and C3 are satisfied.

Let us introduce the single-crossing condition[14] usually considered in a wide range of economic models.

Single-crossing condition. $\left|\frac{\partial \bar{u}}{\partial m} : \frac{\partial \bar{u}}{\partial a}\right|$ is strictly decreasing in *t*.

In the Spence´s model, this condition means that the relative cost of education monotonically decreases with each worker´s ability (type). We can write a strong discrete version of this assumption in the following way:

C4. $u(t, m, a)$ satisfies the strong single-crossing condition if when $t_1 < t_2$: $u(t_1, m_1, r_1) \leq u(t_1, m_2, r_2)$ and $m_1 \leq m_2$ imply $u(t_2, m_1, r_1) \leq u(t_2, m_2, r_2)$ and strictness in either inequality implies $u(t_2, m_1, r_1) < u(t_2, m_2, r_2)$.

C4 coincides with assumption A4 of Cho and Sobel (1990) and it implies that higher types are more willing to choose higher signals than lower types.

Under assumptions C1-C4, it is easy to see that the intuitive criterion will not discard any equilibria. For example, in any of the pooling equilibria obtained, assumption C2 guarantees that all types of sender might benefit from a deviation whenever the receiver assigns a sufficiently high probability to the highest possible type after observing that deviation. Hence, we cannot find a message which is "dominated" in equilibrium for a particular type of sender. The same will occur with a semipooling equilibrium as those considered in theorem 1.

Nevertheless, it is easy to find a refinement suggested by previous literature that selects a unique equilibrium in the class of signaling games in which C1-C4 are satisfied. For example, the divinity criterion would choose a unique equilibrium in which all the types of sender send

[14] Athey (2001) proves the existence of pure strategy Nash equilibria in a generic class of games with incomplete information and characterizes the type of equilibria obtained when the single crossing condition is satisfied in a wide variety of games.

the maximum possible message. Let us explain how a pooling equilibrium in which all types of sender send a lower message than m_N will be discarded. According to C1 and C4, all the receiver´s best responses to a higher message that give a type of sender at least the same utility level than in that pooling equilibrium would make a higher type strictly better off than in the equilibrium. Thus, any equilibria in which all types of sender are sending a message lower than m_N would be discarded by the divinity criterion.

For the same reason, the new criterion suggested in this paper will also select the same pooling equilibrium. Again, let us consider a pooling equilibrium in which all types of sender send a message which is lower than m_N and imagine that a *t-type* of sender deviates from that pooling equilibrium by sending a higher message. According to C1 and C4, the higher the type of sender, the greater the maximum profit he could make from that deviation. Thus, after observing m_N, which is higher than that in the pooling equilibrium considered, the receiver should assign a greater posterior probability to the highest possible type of sender. In such a case, C2 implies that the t_n*-type* of sender would prefer to deviate from the equilibrium. In this way, we rule out all the pooling equilibria in which all types of sender send a message lower than m_N. However, if all the types of sender are sending the maximum possible message in equilibrium, a deviation to a higher message is not possible. Therefore, the only equilibrium that will survive our refinement is also the pooling equilibrium in which all types of sender send the maximum possible message.

In some contexts, such as the example represented in figure 1, C4 is not satisfied. In order to prove that C4 is not satisfied in this game, the wimp will be denoted by t_1, whereas the surly will be represented by t_2. Moreover, r_1 will represent a pure strategy consisting of not dueling whatever the message observed by the receiver, while r_2 will be a mixed strategy consisting of assigning the same probability to dueling and not dueling after beer and not dueling after quiche. Finally, m_1 will be quiche and m_2 will be beer. With this notation,

$u(t_1, m_1, r_1) = 3$, because it is the utility obtained by the wimp when he chooses quiche and the receiver responds to this message by not dueling. Likewise, $u(t_1, m_2, r_2) = 3$ because it is the utility obtained by the wimp when he chooses beer and the receiver chooses the mixed strategy previously described. Additionally, if we assign a strictly lower value to quiche than to beer, $m_1 < m_2$. As a result, if we apply C4, it should imply that $u(t_2, m_1, r_1) < u(t_2, m_2, r_2)$, but this is not satisfied in the example of figure 1. In fact, the utility obtained by the surly when he chooses quiche and the receiver does not duel is $u(t_2, m_1, r_1) = 2$, whereas the utility obtained by the surly when he chooses beer and the receiver chooses the above mixed strategy is $u(t_2, m_2, r_2) = 2$.

As shown in section 3, previous refinements do not allow us to choose a unique equilibrium in the class of signaling games in which C4 is not satisfied. For this reason, we will consider weaker assumption:

C4´: If we denote $\bar{\bar{u}}(m) = \max_{r(a|m) \in MBR(T,m)} \bar{u}(t, m, r(a|m))$, $\left|\bar{\bar{u}}(m_i) - \bar{\bar{u}}(m_j)\right|$ is strictly decreasing in *t*.

The continuous version of this condition will be: $\left|\frac{\partial \bar{\bar{u}}}{\partial m}\right|$ is strictly decreasing in *t*. C4´ means that the cost of the signal decreases with the sender´s type, whenever the receiver chooses the mixed best response which is most preferred by each type of sender. This assumption is weaker than the strong single-crossing condition. In fact, the single-crossing condition requires that the cost of the signal decrease with the sender´s type for all the receiver´s responses to each message, but C4´ only requires that the inverse relationship between the cost of the signal and the sender´s type be satisfied when the receiver chooses the most preferred action by each type of sender.

To end this section, we obtain our final result.

THEOREM 2. *In a signaling game in which C1, C2, C3 and C4´ are satisfied, a unique equilibrium will survive the refinement of the most profitable deviator. The resulting equilibrium will be a pooling equilibrium in which all types of sender choose the highest possible message.*

Proof. Imagine that a type of sender is sending a message and given that message, the receiver is choosing the strategy which gives this type of sender the highest possible utility level. If a greater message gives this type of sender a greater utility level in his best possible scenario, C4´ means that this particular profit made from the additional message will increase with the type of sender. Suppose we have a pooling equilibrium in which all the types of sender send the same message, but it is lower than m_N. Now imagine that a type of sender deviates from that equilibrium by sending m_N. C4´ implies that the highest type of sender would make a greater profit from this deviation, when the receiver chooses the mixed best response to that deviation which is most preferred by each type of sender. Thus, according to our refinement, the receiver should assign a greater probability to the highest possible type of sender after observing such a deviation from the equilibrium. Consequently, C2 implies that any type of sender would want to deviate from the pooling equilibrium considered given those posterior probabilities assigned after the deviation. Using a similar argument, we could discard all the semipooling equilibria considered in theorem 1. Obviously, none of the types can send a higher message when they are already sending the highest possible message. Thus, the pooling equilibrium in which all the types of sender are sending the highest possible message will survive our refinement.

However, the refinements developed by previous literature will not necessarily select a unique equilibrium in signaling games in which C1, C2, C3 and C4´ are satisfied. In fact, the game in figure 1 provides an example of a signaling game satisfying all those assumptions, but none of the previous refinements can discard any of the pooling equilibria obtained.

8. Conclusions

The main purpose of this research was to find a refinement that selects a unique pooling equilibrium in a generic class of signaling games. However, our primary goal was not to propose an equilibrium refinement that selects a unique equilibrium in every game. In fact, there might be some games in which no equilibria survive our refinement. Despite this potential problem, we have proven that our refinement is useful to select a unique equilibrium or to restrict the number of equilibria considered in some contexts where all the previous refinements were useless.

Interestingly, only a weaker version of the single-crossing condition will be required in order to select a unique equilibrium in a monotonic signaling game using our criterion. The key assumptions in the generic class of signaling games we studied in this article were a limited available set of signals and a sufficiently low cost of the signals. As a result of those assumptions, multiple pooling equilibria survived the refinements developed by previous literature. Therefore, our refinement will be useful in those contexts in which those key assumptions are satisfied. For example, let us consider the well-known signaling game introduced by Spence (1973) to analyze the role of education to signal workers´ ability. In this model, education will be an effective signal of ability whenever the cost of education decreases sufficiently with workers´ ability and the maximum possible level of education is sufficiently high. These assumptions may not be satisfied in many real situations. For example, in some countries the monetary cost of college is low and it is not so clear that the psychological cost of getting a diploma decreases sufficiently with workers´ ability.

In finance, authors have extensively used signaling games and our refinement might also be especially useful because the rate at which the cost of signaling declines with the sender´s type is not sufficiently high or because the number of signals the sender can send is sufficiently limited. For example, Leland and Pyle (1977) present a signaling model to explain the proportion of the shares of a firm bought by its managers. If managers are risk averse, it will be costly to commit to holding a sizeable fraction of their portfolios in the firm, rather than be fully diversified. Once again, there is a unique separating equilibrium in which the fraction of shares held by the managers of a firm signals its future expected profit to potential investors. Nevertheless, the choice of this unique equilibrium hinges on the assumptions that each manager can hold a sufficiently high number of shares of the firm and the cost of holding more shares is sufficiently higher for insiders who have a low quality firm. However, there are laws that limit the quantity of shares that each manager can hold in order to prevent managers from taking advantage of insider information, in which case the maximum level of the signal is limited. In this context, there is likely to be multiple pooling equilibria that will survive all the refinements developed by previous literature.

This refinement may be easily extended to cheap talk games. In typical signaling models, there are exogenous differences in signaling costs. However, in typical cheap talk games, what creates the endogenous signaling costs is the receiver´s choice of action. In this type of games, sending certain message will be costly if the receiver´s best response to that signal makes the sender worse off than sending an alternative message. Now, the sender´s cost of a message may be measured as the drop in the sender´s payoff when sending that message in comparison to his payoff with an alternative message. Once again, if the set of actions that the receiver may choose are limited, previous refinements might lead to multiple equilibria. In those games, it would be easy to adapt assumptions C1, C2, C3 and C4´ in order to obtain uniqueness. It would suffice to substitute the assumptions about the sender´s preferences for messages with similar assumptions about the sender´s preferences for the receiver´s actions.

References

Athey, S. (2001): "Single crossing properties and the existence of pure strategy equilibria in games of incomplete information", *Econometrica*, 69(4), pp. 861-889.

Banks, J. S. and Sobel, J. (1987): "Equilibrium selection in signalling games", *Econometrica*, 55(3), pp. 647-661.

Bhattacharya, S. (1979): "Imperfect information, dividend policy and "the bird in the hand" fallacy", *Bell Journal of Economics*, 10(1), pp. 259-270.

Cho, I-K. and Kreps, D. M. (1987): "Signaling games and stable equilibria", *Quarterly Journal of Economics*, 102(2), pp. 179-202.

Cho, I-K. and Sobel, J. (1990): "Strategic stability and uniqueness in signalling games", *Journal of Economic Theory*, 50, pp. 381-413.

Esó, P. and Schummer, J. (2009): "Credible deviations from signalling equilibria", *International Journal of Game Theory*, 38, pp. 411-430.

Farrell, J. (1993): "Meaning and credibility in cheap-talk games", *Games and Economic Behavior*, 5(4), pp. 514-531.

Fudenberg, D. and He, K. (2017): "Type-compatible equilibria in signalling games", retrieved from https://pdfs.semanticscholar.org/6fdb/83b03341e89dfac9da93eb025c2d9e3bfb20.pdf on the 5th of September, 2018.

Fudenberg, D. and He, K. (2018): "Learning and type compatibility in signaling games", *Econometrica*, 86(4), pp. 1215-1255.

Fudenberg, D. and Tirole, J. (1991): *Game Theory*. MIT Press, Cambridge.

Grossman, S. J. and Perry, M. (1986): "Perfect sequential equilibrium", *Journal of Economic Theory*, 39, pp. 97-119.

John, K. and Williams, J. (1985): "Dividends, dilution, and taxes: A signalling equilibrium", *Journal of Finance*, 40(4), pp. 1053-1069.

Kohlberg, E. and Mertens, J. F. (1986): "On the strategic stability of equilibria", *Econometrica*, 54(5), pp. 1003-1037.

Leland, H. E. and Pyle, D. H. (1977): "Informational asymmetries, financial structure, and financial intermediation", *Journal of Finance*, 32(2), pp. 371-387.

Mailath, G. J., Okuno-Fujiwara, M. and Postlewaite, A. (1993): "Belief-based refinements in signaling games", *Journal of Economic Theory*, 60, pp. 241-276.

McLennan, A. (1985): "Justifiable beliefs in sequential equilibrium", *Econometrica*, 53, pp. 889-904.

Milgrom, P. and Roberts, J. (1986): "Price and advertising signals of product quality", *Journal of Political Economy*, XCIV, pp. 796-821.

Riley, J. G. (1979): "Informational equilibrium", *Econometrica*, XLVII, pp. 331-59.

Riley, J. G. (2001): "Silver signals: Twenty-five years of screening and signaling", *Journal of Economic Literature*, XXXIX, pp. 432-478.

Sobel, J., Stole, L. and Zapater, I. (1990): "Fixed-equilibrium rationalizability in signalling games", *Journal of Economic Theory*, 52, pp. 304-331.

Spence, A. M. (1973): "”Job market signalling", *Quarterly Journal of Economics*, 87(3), pp. 355-379.

Spence, A. M. (1974): *Market Signaling*. Cambridge, MA: Harvard University Press.