Journal of Economics
https://doi.org/10.1007/s00712-024-00882-x

Signaling games with a highly effective signal

Miguel Ángel Ropero García[1]

Received: 16 February 2024 / Accepted: 17 August 2024

Abstract
We study a class of signaling games in which one of the signals induces the receiver to take an action that provides the sender with the highest utility. This class of games has multiple pooling equilibria, but the equilibrium in which all senders′ types choose the signal that induces the receiver to take that action is more plausible than others. Although all the equilibria in pure strategies are divine in our class of games when the single-crossing condition is not satisfied, only the plausible equilibrium is a neologism-proof equilibrium. Therefore, we have identified a general class of signaling games in which the neologism-proof equilibrium is useful to select the most plausible equilibrium, whereas all the pooling equilibria survive divinity and other less restrictive refinements. We apply our model to an educational signaling game with two features. First, the highest level of education allows a worker to access a more productive segment of the labor market. Second, the educational system is non-selective and consequently, the cost of education does not change with the worker′s ability. As expected, there is overeducation in equilibrium because all worker′s types choose the highest level of education.

Keywords Education · Monotonic game · Neologism-proof equilibrium · Signaling game · Single-crossing condition

JEL Classification C72 · C73 · C79 · D82

1 Introduction

In standard signaling games, a sender has private information and a receiver tries to infer that information from a message sent by the sender. There are some environments in which there is a message that induces the receiver to take an action that maximizes the sender′s payoff. For example, Kihlstrom and Riordan (1984) showed

✉ Miguel Ángel Ropero García
maropero@uma.es

[1] Department of Applied Economics, Faculty of Economics and Business, University of Malaga, El Ejido, 6, 29071 Malaga, Spain

Published online: 04 September 2024

that firms may use costly advertising in order to signal product quality, but in the settings we have in mind, a promotional campaign might also make more consumers aware of the product and consequently, that campaign may be more beneficial to companies than no advertising.[1] Similarly, Spence (1973) suggested that workers invest in education in order to signal their productivity in the labor market, but a higher level of education may also give access to better job opportunities in the situations we consider, in which case education may be beneficial to the worker. In both examples, the sender prefers the receiver′s response to one of the messages. Specifically, some consumers may consider buying a product after observing a marketing campaign that makes them aware of its existence, but that product would be overlooked by those consumers without the advertising expenditure. In the second example, a company may consider offering a highly productive job only to those workers whose level of education is sufficiently high. Finally, it is easy to imagine situations in which only the highest sender′s type would have access to the receiver′s best responses with perfect information. For example, with perfect information, those customers with the highest willingness to pay for a product would prefer to buy the product from only the best companies. Likewise, companies would prefer to offer certain high-profile jobs to only the most productive workers.

In this article, we analyze a general class of signaling games in which the highest message induces the uninformed receiver to take the action that is most preferred by the sender. Additionally, the receiver′s best response to the highest type of sender with perfect information is so good that all sender′s types will have strong incentives to imitate the highest. In this type of games, there will only be pooling equilibria and as expected, the outcome in which all sender′s types choose the highest message is the only equilibrium selected by the neologism-proof refinement developed by Farrell (1993). However, other standard refinements do not rule out the implausible equilibria.

As out-of-equilibrium beliefs are unconstrained in a perfect Bayesian equilibrium, there are usually multiple equilibria in signaling games and previous researchers have proposed different refinements in order to eliminate some of those equilibria. In the literature on game theory, there is a wide variety of articles which have identified a general class of signaling games in which the unique equilibrium selected by previous refinements is a separating equilibrium (Riley 1979; Cho and Sobel 1990; Esó and Schummer 2009; Liu and Pei 2020). Differently, Mailath et al. (1993) use a numerical example of an educational signaling game in order to illustrate that the unique result selected by their refinement will be a separating or a pooling equilibrium depending on the prior distribution of worker′s types. Unlike that literature, we identify a general class of signaling games in which there are only

[1] According to the taxonomy introduced by Johnson and Myatt (2006), an advertisement consists of both hype and real information. The hype corresponds to basic publicity for a product from which a consumer may learn of a product existence, availability, price and any objective quality, that is, a characteristic that is valued by every consumer. Therefore, hype will always increase demand. In contrast, real information allows consumers to evaluate their subjective preferences for a product. Hence, real information increases the dispersion of consumers′ valuations and rotates the demand curve for a product. We are considering companies′ use of advertising as a signal when advertisements include hype.

pooling equilibria, but the refinement introduced by Farrell (1993) selects the plausible equilibrium. Interestingly, divinity or D1 as defined by Cho and Kreps (1987) and Cho and Sobel (1990) would also select the same equilibrium in our class of games if we added the single-crossing condition, but we found that Farrell′s criterion does not need that condition to select the same equilibrium.

As an application, we study an educational model which belongs to our general class of signaling games. Here, the worker′s productivity in the labor market increases with his level of education. Additionally, only those workers with the highest level of education will gain access to a highly productive job with training opportunities. Furthermore, the productivity difference between worker′s types is so high for all levels of education that low-productivity workers will have strong incentives to imitate high-productivity workers′ educational investment. Finally, the prior distribution of worker′s types is sufficiently concentrated on the highest possible type. Unsurprisingly, in a non-selective educational system with low costs of education, we obtain a unique neologism-proof equilibrium in pure strategies in which all worker′s types choose the highest possible level of education in order to get the job with training opportunities. However, divinity and other criteria do not eliminate the other implausible equilibria.

This article is organized as follows. Section 2 briefly reviews previous literature and shows our main contribution. Section 3 presents the main ingredients of a typical signaling game, whereas Sect. 4 introduces the main assumptions which differentiate our general class of games from other signaling games and characterizes all the equilibria obtained. After this, Sect. 5 proves that the refinement proposed by Farrell (1993) selects the most plausible equilibria even when the single-crossing condition is not satisfied, whereas other refinements require that condition in order to select that equilibria. Next, in Sect. 6, we study a similar model to that introduced by Spence (1973) in his seminal article as an application of our class of games. Finally, Sect. 7 summarizes the main conclusions and the appendix includes the proofs of all lemmas, propositions and theorems.

2 Literature review and contribution

As there are multiple equilibria in most signaling games, previous literature has proposed different criteria in order to constrain the out-of-equilibrium beliefs and in many cases, the unique equilibrium selected is a separating equilibrium (Riley 1979; Cho and Sobel 1990; Esó and Schummer 2009; Clark and Fudenberg 2021). Similarly, Mailath et al. (1993) use a specific example of the Spence′s model with separating and pooling equilibria in order to illustrate that their criterion only selects the least-cost separating equilibrium under certain assumptions, but they also found the conditions under which the selected equilibrium is pooling. Unlike that literature, we study a general class of signaling games in which there are only pooling equilibria, but one of them is more plausible than others. In those games, we show that only the refinement introduced by Farrell (1993) will be useful to select that plausible equilibrium.

In his seminal article, Farrell (1993) defined the neologism-proof equilibrium (NPE), which is a sequential equilibrium with no credible neologism. A credible neologism is an out-of-equilibrium message for which there is a set of sender′s types who are strictly better off than in equilibrium when the receiver responds to that deviation by concentrating his beliefs on those types, whereas other types are worse off. Farrell (1993) called pooling equilibria uncommunicative and he proved that the NPE are uncommunicative when the sender′s preferences over the receiver′s beliefs are independent of the sender′s type. However, Farrell (1993) acknowledged that the NPE needs not be unique if it exists. In this article, we have identified a general class of monotonic signaling games in which there are multiple pooling equilibria that survive standard refinements, but only one is NPE. Therefore, the concept of equilibrium introduced by Farrell is more useful than other refinements in our class of games.[2]

Interestingly, Cho and Sobel (1990) found that their refinement, namely divinity, selects a unique pooling equilibrium in a class of signaling games with the single-crossing condition. In our class of games, we found that the set of divine equilibria coincides with the set of NPE when we add that condition. As Farrell′s criterion does not need this assumption in order to select those equilibria, it is more robust to the specification of the single-crossing condition than divinity in our specific class of signaling games.

In comparing different refinements, there is an abundance of literature demonstrating that divinity is the strongest criterion. For example, Cho and Kreps (1987) showed that the intuitive refinement gives rise to a lower set of equilibria than those criteria suggested by McLennan (1985), Milgrom and Roberts (1986) and Kohlberg and Mertens (1986) and that all divine equilibria also satisfy the intuitive criterion, but the opposite is not true. Furthermore, Sobel et al. (1990) proved that the set of equilibria that survives divinity is a subset of a set of equilibria obtained after deleting iteratively weakly dominated strategies. Additionally, Cho and Sobel (1990) showed that divinity, universal divinity and never a weak best response criterion are equivalent in monotonic signaling games. Therefore, divine equilibria are strategically stable as defined by Kohlberg and Mertens (1986) in monotonic signaling games. Finally, Fudenberg and He (2017, 2018, 2020) proved that the set of divine equilibria is a subset of the set of type compatible equilibria and of the set of rationality compatible equilibria, Clark and Fudenberg (2021) showed that justified communication equilibria are path equivalent to divine equilibria in co-monotonic

[2] In a similar vein, Grossman and Perry (1986) defined a perfect sequential equilibrium (PSE) in which the sender′s strategy is a best response to the receiver′s strategy and the receiver′s strategy is a best response to the sender′s strategy and to *credible beliefs*. The concepts introduced by Grossman and Perry and Farrell are closely related because the receiver′s *credible beliefs* after an out-of-equilibrium message should assign positive probabilities to those sender′s types who can be better off than in equilibrium. As a result, if there is an out-of-equilibrium message and there are no *credible beliefs* that support an equilibrium, both refinements rule out that equilibrium. Then, every NPE is a PSE, but the opposite is not true (see the example of Fig. 5 in Grossman and Perry (1986)). Finally, as Grossman and Perry (1986) showed that the set of PSE is a subset of the set of intuitive equilibria (IE), we conclude that $NPE \subseteq PSE \subseteq IE$.

signaling games, and Dominiak and Lee (2023) proved that each divine equilibrium is a rational hypothesis testing equilibrium, which is their notion of equilibrium. As we consider monotonic signaling games and those articles showed that divinity selects a subset of the set of equilibria selected by those refinements mentioned above in monotonic signaling games, we only compare the neologism-proof equilibrium selected to that selected by divinity with and without the single-crossing condition.

Before applying any refinement, we have identified a class of games in which all the equilibria in pure strategies are pooling, which are monotone equilibria as defined by Liu and Pei (2020) because lower sender′s types do not choose higher messages than higher types. Our model and that introduced by Liu and Pei (2020) complement each other because both articles provide sufficient conditions under which all equilibria in pure strategies are monotone. They require the sender′s utility function to be strictly decreasing with the sender′s message, whereas cheap-talk games in which the sender′s utility function does not depend on the message may also belong to our class of games.[3] Additionally, Liu and Pei assume that the sender′s utility function is strictly increasing in the receiver′s action, while we only require that function to be weakly increasing. Furthermore, their receiver′s payoff function satisfies a quasiconcavity-preserving property, but we assume that it has increasing differences in the sender′s type and the receiver′s actions. The main difference between our models is that they assume that the sender′s utility function has strictly increasing differences in her types and messages and it has increasing differences in her types and the receiver′s actions, but we do not even need a single-crossing condition. However, Liu and Pei do not make assumptions about the prior distribution of sender′s types, whereas we assume that it is sufficiently concentrated on the highest possible type. As a result of that assumption, all sender′s types have strong incentives to imitate the highest type in our model.

Our article shares some similarities with that by Pei (2020). For example, a key assumption of our general class of signaling games is that the receiver′s best response to the sender′s highest message depends on the sender′s type. This is also one of the assumptions of the model introduced by Pei (2020). Additionally, in his theorem 2, he also assumes that the receiver′s highest action best replies against the sender′s highest type and message. This is also a key assumption in our general class of signaling games that gives strong incentives to all the sender′s types to imitate the highest when he sends the highest possible message. Apart from those key assumptions, like Pei, we assume that the receiver′s best response to each message is a singleton under perfect information, which is satisfied for generic payoff functions of the receiver. However, there are important differences between our class of games and that analyzed by Pei (2020). First, he assumes that the sender′s utility function is strictly decreasing with the message, but our class of signaling games also accommodates cheap-talk games. Secondly, he considers sender′s utility functions that are strictly increasing with the receiver′s action, but we allow those utility functions to

[3] See a simple application of our general model to a cheap-talk game about the interaction between a child and his mother in the online appendix.

be weakly increasing. A key difference between both models is that the sender′s utility function has increasing differences in the sender′s type and messages and actions in Pei′s model. Although we do not impose this type of single-crossing condition, all the sender′s types have strong incentives to imitate the highest possible type in our specification of the sender′s utility function. Finally, Pei considers signaling games in which the receiver′s utility function has strictly increasing differences in his action and the sender′s types and messages, whereas we assume weakly increasing differences. Therefore, our models can be seen as complements because Pei studies reputation building in repeated signaling games, whereas we consider one-shot signaling games with a different type of monotonicity condition.

In this article, we also study an extension of the model introduced by Spence (1973) that satisfies all the assumptions of our class of signaling games. In particular, we assume that education contributes to increasing productivity in the labor market. In our setting, a worker has private information on his productivity and chooses an education level. After that, a company observes the worker′s decision and chooses the type of job offered and the wage associated to that job. If the worker′s level of education is different from the maximum level, the worker′s productivity inferred by the company is so low, that only a low-quality job without training opportunities can be offered and the wage paid will equal the expected productivity. However, if the worker achieves the maximum level of education, the company may find it profitable to offer a high-quality job with training opportunities and a higher wage. As the educational system is not selective in the model, we obtain a plethora of pooling equilibria, but the only NPE in pure strategies is that in which the highest level of education is selected. Interestingly, Dominiak and Lee (2023) use their refinement to constrain the set of pooling equilibria in the standard Spence′s model, but they select equilibria in which the level of education chosen by the worker is lower than a certain threshold. Additionally, they cannot discard the Riley′s separating equilibrium with their criterion.

3 Model

Now, we can introduce the main elements of our generic class of signaling games. 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, $T = \{t_1, t_2, \cdots, t_n\} \subset \mathbb{R}$, where $t_1 < t_2 < \cdots < t_n$. 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 $m \in M = \{m_1, m_2, \cdots m_N\}$ and $m_1 < \cdots < m_N$. After observing *m*, *R* chooses an action, *a*, from a finite partially ordered set *A(m)* and both players obtain their levels of utility. The sender and the receiver have von Neumann-Morgenstern utility functions *u(t, m, a)* and *v(t, m, a)*, respectively. As usual, we assume that u(·) increases with *a*. Furthermore, the receiver′s utility function has increasing differences in *t*, that is, when $a' > a$ and $t' > t$, $v(t', m, a') - v(t', m, a) \geq v(t, m, a') - v(t, m, a)$ for all m. This assumption means that the receiver′s marginal utility from his action does not decrease with the sender′s type.[4]

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

As usual, we consider perfect Bayesian equilibria (PBE), which are sets of signaling rules, $q(m|t)$, action rules, $r(a|m)$, and beliefs, $\mu(t|m)$, that satisfy two conditions:

1. Rationality: Each player maximizes his or her utility given the other player′s strategy:

   i. $\forall t \in T, q(m^*|t) > 0$ only if $\overline{u}[t, m^*, r(a|m^*)] = \max_{m \in M} \overline{u}[t, m, r(a|m)]$.
   ii. $\forall m \in M, r(a^*|m) > 0$ 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 $\overline{u}[t, m, r(a|m)] = \sum_{a\in A(m)} u(t, m, a)r(a|m)$.
2. Consistency: The equilibrium posterior beliefs are consistent with Bayes′ rule: If $\sum_{t\in T} q(m|t)p(t) > 0$, then $\mu(t^*|m) = \frac{q(m|t^*)p(t^*)}{\sum_{t\in T} q(m|t)p(t)}$.

From now on, $\overline{u}(\cdot)$ $(u(\cdot))$ denotes S′s utility when R chooses a mixed (pure) strategy.

4 Potential equilibria

In this section, we characterize our class of games and the equilibria obtained. To start with, we introduce some notation: Let $BR(I, m)$ represent the union of sets of mixed best responses to assessments concentrated on the subset I of T and similarly, $BR_{PB}(I, m)$ is the set of mixed best responses to m when all the sender′s types belonging to I send m with probability one, but the rest of types send it with probability zero, and the receiver′s posterior beliefs after m are prescribed by Bayes′ rule in that situation. Finally, $a(t, m) = arg \max_{a\in A(m)} v(t, m, a)$ represents R′s optimal response

[4] This assumption is satisfied in many applied signalling games. For example, in the model proposed by Spence (1973), employers′ marginal utility from a greater wage does not decrease with the worker′s productivity because an employer′s utility from the wage offered to a worker is equal to her productivity minus the wage paid. Similarly, in the limit pricing model introduced by Milgrom and Roberts (1982), the potential entrant′s profit from entering the market increases with the incumbent′s cost.

to m when he knows that S's type is t and $a_p(m) = arg \max_{a \in A(m)} \sum_{t \in T} v(t, m, a)p(t)$ denotes R′s optimal response to m under his prior beliefs.

We assume that all those best responses exist for each value of m and for each set of posterior beliefs. Additionally, for each t and each m, $a(t, m)$ and $a_p(m)$ are unique pure strategies and due to increasing differences of $v(\cdot)$, $a(t, m)$ is non-decreasing with t and $a(t_1, m) \leq a_p(m) \leq a(t_n, m) \forall m \in M$.

Besides, we assume that $a_p(m_N) = a(t_n, m_N)$. According to this assumption,[5] when the posterior beliefs are equal to the prior probabilities, R will choose the same response to the highest possible message as that chosen when R is certain that the sender has the highest possible type. This assumption may be satisfied in two different scenarios. First, when the prior probability of the highest possible type is sufficiently high, the receiver will respond to a message in a pooling equilibrium as if he were playing against that type. Second, after observing m_N, the receiver may be much more penalized from not choosing his optimal action against t_n than from not choosing his optimal action against other sender′s types. In this second scenario, the receiver has also strong incentives to respond to m_N by choosing his best response to t_n. In our class of games, an equilibrium in which all sender′s types send the highest message will be more plausible than others due to this specific assumption.[6]

Since we are assuming that $u(\cdot)$ increases with a, our class of signaling games are monotonic as defined by Cho and Sobel (1990) because all types of sender have the same preferences for the receiver′s strategies. The next assumptions characterize the subset of monotonic signaling games we study:

C1 The following conditions are satisfied:

i. $u(t, m_N, a(t_n, m_N)) > u(t, m, a(t_n, m)) \forall t \in T \quad \forall m \neq m_N$
ii. For all $I, J \subset T$ such that $I \cap J = \emptyset$ and $t_n \in I$, for all $m, m' \in M$, with $m \neq m'$, for all $r_m \in BR_{PB}(I, m)$ and for all $r_{m'} \in BR_{PB}(J, m')$, then there will be some $t \in J$ such that $\overline{u}(t, m, r_m) > \overline{u}(t, m', r_{m'})$.

[5] This assumption is the same as assumption 4 in Pei (2020). Pei refers to a prior distribution function that satisfies this assumption as an optimistic distribution. As shown by Pei, this is a distribution that attaches high enough probability to the highest sender′s type. Pei (2020) also shows the conditions under which optimistic posterior beliefs arise in a repeated signalling game with reputation concerns.

[6] Let us use a simple example to illustrate this assumption in the animal kingdom. For instance, Flower (2011) shows that fork-tailed drongos use deceptive alarm calls to steal food from meerkats. In this example, the sender is a drongo who may have observed a predator (high type) or not (low type). The drongo may use an alarm call (high message, m_N) or not (low message). The meerkat is the receiver who may continue handling food or flee to cover. In this case, the outcome described by Flower in which drongos use false alarms is a pooling equilibrium in which all drongos′ types, those who observe a predator and those who do not, use the alarm call (m_N). In this example, meerkats response to that alarm by fleeing to cover, which is the same response to the alarm when they know that drongos watched a predator, which implies that $a_p(m_N) = a(t_n, m_N)$. As shown above, this may happen for two reasons. First, when there are many predators, the probability of observing one will be high and meerkats will response to alarms by fleeing. Second, meerkats will be killed if they do not flee to cover when there is a predator and for this reason, they have strong incentives to flee just in case.

C.2 For all $t \in T$ and for all $m, m' \in M$: $u\big(t, m', a\big(t_1, m'\big)\big) \leq u\big(t, m, a_p(m)\big)$.

The first part of C1 means that all types of sender prefer the highest possible message to a different message when the receiver knows that those messages were sent by the highest type of sender and responds to them accordingly.[7] This assumption is satisfied when the highest message sent by the highest type makes the receiver choose some action that would never be chosen with a different message. For example, imagine that education is productive, and is especially productive among those workers with the highest ability in Spence′s model. In this scenario, if employers responded to each level of education by paying the monetary value of the productivity of the highest-ability worker with that level of education, all workers would choose the highest level of education regardless of their ability. According to part ii in C1, any type of sender will always prefer to send a different message whenever this new message makes the receiver believe that he or she belongs to a set of types with the highest possible type. This assumption implies that the receiver′s response to the highest possible type with certainty is so good that all types will try to imitate him. In our educational example, if the productivity of the worker with the highest ability is much greater than the productivity of workers with different abilities, the former will receive an extremely high wage when he is identified by employers and for this reason, workers with different abilities will have strong incentives to mimic the highest-ability worker by choosing the same level of education. According to C2, any type of sender will be weakly better off with a message when the receiver reacts to that message by using the prior beliefs than with any message when the receiver responds to it by concentrating his beliefs on the lowest type. This is a standard assumption in many signaling games. In our educational example, it means that a worker prefers one level of education to another when employers respond to the former with a wage equal to the ex-ante expected productivity and responds to the latter with a wage equal to the productivity of the lowest-ability worker.[8] In this educational example, assumptions C1 and C2 are satisfied as long as the cost of education is sufficiently low regardless of the worker′s ability. Otherwise, low-ability workers would find it prohibitively expensive to choose the same level of education as the highest-ability worker.

Our assumptions describe situations in which the cost of the signal is sufficiently low and the receiver′s response to the best possible sender is so good that all sender′s types have strong incentives to imitate the highest type. This is the reason why there are multiple pooling equilibria. Additionally, as there is one message that makes the receiver respond to the highest type of sender by choosing the sender′s favorite action, we should anticipate a result in which the highest sender chooses that message and other types make the same choice. Consequently, a pooling equilibrium in

---

[7] If the preferred message is different from the highest, we rename the messages and the assumption still holds.

[8] Interestingly, Pei (2020) includes a numerical example of a one-shot signaling game that satisfies all of our assumptions (see page 2176 in that article), but his main purpose is analyzing repeated signaling games.

which all sender's types send the highest possible message will be more plausible than others under the assumptions of the model.

Since the game is finite, the existence of equilibrium is guaranteed. Now, we can characterize the equilibria in pure strategies obtained in our class of games.

Lemma 1 *Under assumptions C1 and C2, there are only N PBE in pure strategies, which are completely pooling and in each PBE, all sender's types send the same message from* $M = \{m_1, m_2, \cdots, m_N\}$.

Under assumption C2, any outcome with all types sending the same message is an equilibrium because there is a response to any other message that prevents S from defecting. Part ii of assumption C1 guarantees that the sender will always find it profitable to convince the receiver that he belongs to the set with the highest type. For this reason, a type belonging to a set without t_n would prefer to deviate from a non-pooling equilibrium by sending the message sent by the set with t_n.

Now, in order to identify the PBE in mixed strategies obtained, we introduce two definitions.

Definition 1 We say that *R's beliefs* are *induced by prior probabilities concentrated on I* whenever they are defined as: $p_I(t) = \frac{p(t)}{\sum_{t\in I} p(t)}$ if $t \in I$ and $p_I(t) = 0$ otherwise. Let us call $P_I(t_i) = \sum_{t=t_1}^{t=t_i} p_I(t)$, which is the cumulative distribution function associated with $p_I(t)$.

Imagine that there were a PBE in which only those types belonging to I choose m with probability one. In that equilibrium, R's beliefs after m would be calculated using $p_I(t)$ and R would respond to m with a best response from $BR_{PB}(I, m)$.

Definition 2 If only a set $I \subset T$ sends m with positive probability in a PBE, we say that *R's beliefs are induced by that equilibrium message* whenever they are defined as: $\mu_I(t|m) = \frac{q(m|t)p(t)}{\sum_{t\in I} q(m|t)p(t)}$ if $t \in I$ and $\mu_I(t|m) = 0$ otherwise. Similarly, $M_I(t_i|m) = \sum_{t=t_1}^{t=t_i} \mu_I(t|m)$.

In a PBE in which only those types belonging to I send a message, definition 2 implies that R's posterior beliefs after that message will be prescribed by Bayes′ rule.

Now, we use those definitions in order to limit the PBE in mixed strategies obtained.

Lemma 2 *In a signaling game in which C1 holds, there cannot be an equilibrium in mixed strategies which satisfies the following conditions simultaneously:*

i. A set of sender's types, $I \subset T$, *send a message,* $m \in M$, *and another set,* $J \subset T$ *, send another message,* $m' \in M$ *, with positive probabilities, where* $t_n \in I$, $I \cap J = \varnothing$ *and* $m \neq m'$.

ii. *The receiver's beliefs induced by m, $M_I(t|m)$, first-order stochastically dominate the beliefs induced by the prior probabilities concentrated on I, $P_I(t)$.*
iii. *The receiver's beliefs induced by the prior probabilities concentrated on J, $P_J(t)$, first-order stochastically dominate the beliefs induced by m′, $M_J\left(t|m'\right)$.*

This lemma allows us to restrict the number of equilibria in mixed strategies significantly. According to this result, any sender′s type will have incentives to send the same message as that sent by the highest type whenever that message increases the type expected by the receiver. In fact, part ii of C1 gives any sender′s type strong incentives to imitate the highest type because the receiver′s response to that type provides the action preferred by all sender′s types.

The next result is a summary of lemmas 1 and 2.

Theorem 1 *In a signaling game in which C1 and C2 are satisfied, there will only be two types of PBE. First, there are N completely pooling PBE in pure strategies and in each equilibrium, all types send one of the messages from $M = \left\{m_1, m_2, \cdots, m_N\right\}$ with probability one. Second, there will be equilibria in mixed strategies which do not satisfy the conditions of lemma 2.*

This theorem shows that there are only completely pooling PBE in pure strategies. For example, imagine that we have a signaling game with $T = \{1,2,3,4,5,6,7,8,9,10\}$, $M = \left\{m_1, m_2, m_3\right\}$ and satisfying assumptions C1 and C2. In this game, there will be 3 pooling PBE in pure strategies: one in which all types send m_1, another in which all types send m_2 and the last in which all types send m_3. Additionally, there is a multiplicity of PBE in mixed strategies. However, lemma 2 discards some potential equilibria. For instance, there cannot be an equilibrium in which *type 10* chooses m_1 and m_3 with positive probability, *types 1–4* and *6–9* choose m_3 with probability one, and *type 5* chooses m_2 and m_3 with positive probability. This equilibrium would satisfy all the conditions of lemma 2. Using the notation of that lemma, $m = m_1$, $m' = m_2$, $I = \{10\}$, $J = \{5\}$ and consequently, condition i of lemma 2 holds. After observing m_1, the receiver′s beliefs should be concentrated on *type 10*. As these beliefs coincide with the beliefs induced by the prior probabilities concentrated on I, the former first-order stochastically dominate the latter and condition ii of lemma 2 is also satisfied. Finally, after observing m_2, the receiver should concentrate his beliefs on *type 5*. As these beliefs are equal to the beliefs induced by the prior probabilities concentrated on J, the latter first-order stochastically dominate the former and condition iii of lemma 2 also holds. Thus, the result in question cannot be a PBE in mixed strategies.

Now, we introduce the single-crossing condition, which is a typical assumption in many applications:

C3 $\overline{u}(\cdot)$ satisfies the single-crossing condition if when $t < t' \ : \quad \overline{u}(t, m, r) \leq \overline{u}\left(t, m', r'\right)$ and $m \leq m'$ imply $\overline{u}\left(t', m', r'\right) - \overline{u}\left(t', m, r\right) \geq \overline{u}\left(t, m', r'\right) - \overline{u}(t, m, r)$ and strictness in either inequality implies $\overline{u}\left(t', m', r'\right) - \overline{u}\left(t', m, r\right) > \overline{u}\left(t, m', r'\right) - \overline{u}(t, m, r)$.

C3 is equivalent to the generalized Spence-Mirrlees condition specified by Liu and Pei (2020), but assumption A4 of Cho and Sobel (1990) provides a different version of this condition. C3 implies that higher types are more willing to choose higher messages than lower types.

5 Equilibrium selection

In this section, we formally introduce the NPE as defined by Farrell (1993) and compare the equilibria selected by this criterion to those selected by divinity.

Definition 3 (Neologism-proof equilibrium). Given a PBE and the utility obtained by the *t-type* of sender in that equilibrium, $\overline{u}^*(t)$, a credible neologism is an out-of-equilibrium message, m, such that there is a set of sender′s types, K, that satisfies the following conditions simultaneously:

$$\overline{u}^*(t) < \overline{u}\left(t, m, r_m\right) \forall t \in K, \forall r_m \in BR_{PB}(K, m) \tag{1}$$

$$\overline{u}^*(t) > \overline{u}\left(t, m, r_m\right) \forall t \in T \setminus K, \forall r_m \in BR_{PB}(K, m) \tag{2}$$

A neologism-proof equilibrium is a PBE that has no credible neologism.

According to this definition, a PBE will pass this test as long as we cannot find a set of types, K, associated with an out-of-equilibrium message, m, that satisfies two conditions. First, when the receiver′s beliefs are induced by the prior probabilities concentrated on K, those sender′s types belonging to K prefer m to their equilibrium message. Second, given those beliefs, every type $t \in T \setminus K$ would prefer the equilibrium message to m.

In order to compare this refinement with the divinity criterion or D1 in Cho and Kreps (1987) and in Cho and Sobel (1990), we include the next definition.

Definition 4 (Divinity) Let (q, r, μ) be an equilibrium and let $\overline{u}^*(t)$ be the equilibrium expected utility of the *t-type* of sender. Given an off-the-equilibrium-path signal, m, we define:

$$P(t|m) = \left\{r \in BR(T, m) \, : \, \overline{u}^*(t) < \overline{u}(t, m, r)\right\} \tag{3}$$

$$P^0(t|m) = \left\{r \in BR(T, m) \, : \, \overline{u}^*(t) = \overline{u}(t, m, r)\right\} \tag{4}$$

where $P(t|m)$ is the R′s set of best responses to m that incentivizes the *t-type* to defect and $P^0(t|m)$ is the set of best responses that gives that type the same utility as in equilibrium.

D1 rules out a disequilibrium message, *m,* of the *t-type* of sender if there exists t' such that:

$$P(t|m) \cup P^0(t|m) \subset P\left(t'|m\right) \tag{5}$$

An equilibrium survives D1 if, for all off-the-equilibrium-path messages m, $\mu(t|m) = 0$ whenever (5) holds for some t' such that $P\left(t'|m\right) \neq \emptyset$.

Now, we show that the set of NPE coincides with the set of equilibria that survives divinity in our class of games when C3 is satisfied.

Proposition 1 *Under assumptions C1-C3, there is only a unique divine equilibrium in pure strategies in which all types of sender send m_N. This is also the unique NPE in pure strategies. Additionally, only those equilibria in mixed strategies in which some sender's type sends m_N with positive probability survive divinity and are NPE.*

It is easy to understand the intuition behind this result. Under part i of C1, when the receiver perfectly identifies the highest type, he responds to m_N with the best possible action. This assumption gives the highest type strong incentives to choose the highest message. Besides, part ii of C1 incentivizes other sender's types to imitate the highest type. For those reasons, the most plausible equilibrium in pure strategies is that in which all types select the highest message and the most plausible equilibria in mixed strategies are those in which some type sends that message. Therefore, when assumptions C1-C3 hold, we select the same equilibria with divinity and with the criterion introduced by Farrell (1993).

As our class of games are monotonic, divinity is equivalent to universal divinity, never a weak best response and not vulnerability to credible deviations as defined by Esó and Schummer (2009). *Thus, Farrell's criterion also selects the same PBE as those refinements in our class of signaling games and the equilibria selected are strategically stable as defined by Kohlberg and Mertens* (1986) *because divinity is equivalent to strategic stability in monotonic signaling games* (*Cho and Sobel* 1990). *Finally, if we substituted assumptions C1 and C2 with a sender's utility function which is strictly quasi-concave, a receiver's utility function with strictly increasing differences and continuous utility functions with respect to m and a, divinity would still select the same equilibrium outcome under the single-crossing condition* (*Cho and Sobel* 1990).

However, our final result shows that we do not need C3 in order to select the pooling equilibrium with the highest message with Farrell's refinement in our class of signaling games, but all equilibria are divine when C3 is not satisfied.

Theorem 2 *Under assumptions C1 and C2, the only NPE in pure strategies is the equilibrium in which all sender's types choose m_N and only those PBE in mixed strategies in which some type chooses m_N with positive probability may be NPE.*

Under our assumptions, there always exists a set of NPE. Specifically, we have proven that this refinement selects the same type of equilibria in our class of signaling games even without the single-crossing condition. In those equilibria, low types imitate the highest type by choosing the highest message. Hence, this theorem shows that we have characterized a general class of signaling games in which the concept *of NPE introduced by Farrell (1993) is more useful than other criteria because it selects the most plausible equilibria in this class of games. Under our assumptions, all sender's types have strong incentives to choose the same message as that chosen by the highest type. Moreover, all sender's types are better off when they choose the highest message as long as the receiver responds to that message by using the prior beliefs. Consequently, the equilibrium in which all sender's types choose the highest possible message is the most plausible result in our class of games and this is the outcome selected by Farrell's criterion.*

6 Economic application: educational signaling

In this section, we analyze an extension of the Spence′s model in which education contributes to increasing productivity and to opening new job opportunities. As the model belongs to our class of signaling games, an overeducation equilibrium will be the only NPE in pure strategies.

In this setting, there is a worker who has private information on his ability. That ability is equal to t_1 with probability p and t_2 with probability $1-p$, where $t_1 < t_2$. Using the usual notation, $T = \{t_1, t_2\}$. After observing her ability, the worker may choose one of the following educational levels: $M = \{e_1, e_2\}$, where $e_1 < e_2$. Let us denote the worker′s cost of each level of education by $c(t, e)$, which increases with the level achieved: $c(t, e_1) < c(t, e_2) \forall t \in T$.

After observing the worker′s level of education, companies will compete for the worker *a la* Betrand and as a result, they pay a wage equal to the worker′s expected productivity. Those companies may offer two types of jobs. First, there are certain inferior jobs that do not require a high level of education and the worker′s productivity in that type of job is given by $\pi_l(t, e_1)$, which increases with her ability, that is, $\pi_l(t_1, e_1) < \pi_l(t_2, e_1)$. In this job, a higher level of education does not increase the worker′s productivity because the skills acquired with that level of education are not related to those required by the job, *i.e.*, $\pi_l(t_1, e_1) = \pi_l(t_1, e_2)$ and $\pi_l(t_2, e_1) = \pi_l(t_2, e_2)$. Besides, the company may offer a second type of jobs that require a minimum level of performance. In particular, the company would be able to sell the products generated by a worker occupying this premium job as long as her productivity is equal to $\overline{\pi}$, that may also measure the quality of the products produced by the worker. For this reason, if the worker′s level of productivity is lower than the required level, the company has to train her so that she can reach it. Before that training, the worker′s productivity in the premium job depends on her ability and educational level and is denoted by $\pi_h(t, e) > \pi_l(t, e) \forall t \in T, \forall e \in M$. Now, this productivity increases with the worker′s ability and with the level of education because some of the mental skills required by the job are acquired with education. Consequently, $\pi_h(t_1, e) < \pi_h(t_2, e) \forall e \in M$ and $\pi_h(t, e_1) < \pi_h(t, e_2) \forall t \in T$. As $\overline{\pi} > \pi_h(t_2, e_2)$, the company should train even the high-ability worker with the highest level of education in the premium job.

Now, we introduce the key assumptions of the model. Firstly, when the worker occupies the premium job, she generates some positive synergy or externality in the company denoted by E and this is only known by the company. In order to simplify the model, we assume that the company′s cost of training in the premium job is linear. Specifically, $K : \mathbb{R} \rightarrow \mathbb{R}^{+}$ is the training cost function defined by:

$$K(x) = \begin{cases} \kappa \cdot x & \text{if} \quad x \geq 0 \\ 0 & \text{if} \quad x < 0 \end{cases} \tag{6}$$

This function means that the training cost generated by increasing the worker′s productivity by $x = \overline{\pi} - \pi > 0$ in the premium job is proportional to that increase in productivity because κ is a constant parameter. In this context, we make the following assumption:

$$\frac{E}{\overline{\pi} - \pi_h(t_2, e_1)} < \kappa < \frac{E}{\overline{\pi} - \pi_h(t_2, e_2)} \tag{7}$$

This assumption implies that the company would only offer the premium job to the high-ability worker with the highest level of education if it had perfect information.[9] Other models, such as those introduced by Gibbons and Waldman (1999) and Liu and Pei (2020), also assume that the worker′s benefit from education increases with her talent.

Next, we add an assumption that guarantees that the company will offer the premium job when it observes the high level of education and is completely uninformed:

$$p < \frac{\frac{E}{\kappa} - \left[\overline{\pi} - \pi_h(t_2, e_2)\right]}{\pi_h(t_2, e_2) - \pi_h(t_1, e_2)} \tag{8}$$

This assumption will be satisfied if the positive externality generated by the worker in the premium job, E, is sufficiently high, or if the prior probability of the low-ability worker is sufficiently low.

When the worker occupies the premium job, she is trained by the company and receives a wage equal to her productivity, $\overline{\pi}$. Although companies compete for hiring the worker, they may differ with respect to the positive synergy generated by the worker in the premium job, in which case the company with the highest synergy may offer the worker a greater wage and will eventually hire the worker. In that case, E measures the difference between the synergy generated by the hiring company and

[9] This assumption is plausible in many contexts. For example, a company may find it prohibitively expensive to provide a worker with the necessary training so that she becomes an intermediate manager who can prepare the account book and financial statement, choose the best financial sources, etc. Even a graduate in business administration may lack some of the necessary skills to work as an intermediate manager just after finishing her education, but providing her with the necessary training will be much less costly.

that generated by the second best, which explains why the worker receives a wage of $\overline{\pi}$ in the premium job after receiving the training.[10]

In this section, we will analyze what happens in a non-selective or universal educational system in which the cost of an additional level of education does not increase with the worker′s ability. Specifically, we assume that $c(t_1, e_2) - c(t_1, e_1) = c(t_2, e_2) - c(t_2, e_1)$. Hence, the single-crossing condition is not satisfied.

In our set-up, $\overline{\pi} > \pi_l(t_2, e_1) + c(t, e_2) - c(t, e_1)$ and consequently, the high level of education is a highly effective signal of productivity because it allows the worker to access a well-paid job with training opportunities. This assumption is satisfied if the company sells the products generated by the worker in the premium job to consumers with a high willingness to pay for the product, or if the training course offered by the company is effective and increases the worker′s productivity in the premium job significantly. In both cases, $\overline{\pi}$ will be sufficiently high and the assumption will hold.

It is easy to see that this model belongs to our general class of signaling games. For example, it is a monotonic game because the worker′s utility function increases with the wage received. Under the assumption described by condition (7), companies never offer the premium job to a worker with the low level of education because her productivity is too low. Therefore, companies will offer the inferior job to a worker with a low or a high level of education with a wage belonging to this closed interval, $[\pi_l(t_1, e_1), \pi_l(t_2, e_2)]$, or they will offer the premium job with a wage equal to $\overline{\pi} > \pi_h(t_2, e_2) > \pi_l(t_2, e_2)$, which implies that companies′ action set is equivalent to offering a wage belonging to $[\pi_l(t_1, e_1), \pi_l(t_2, e_2)] \cup \{\overline{\pi}\}$, which is a partially ordered set with the usual order. The existence of increasing differences in the receiver′s utility function is trivially satisfied because the worker′s productivity increases with her type and for this reason, companies′ profit from paying a low wage to a worker increases with her productivity. As shown above, all the company′s best responses exist and are wages equal to the worker′s expected productivity in the inferior job and $\overline{\pi}$ in the premium job. Similarly, the optimal type of job offered by the company with perfect information is the premium job if the worker has a high ability and a high level of education and the inferior job otherwise. The assumptions shown above guarantee that the company′s best response to the prior beliefs and the low level of education is the inferior job and a wage equal to the expected productivity, $a_p(e_1) = \{\text{inferior job}, p\pi_l(t_1, e_1) + (1-p)\pi_l(t_2, e_1)\}$, whereas the best response to the prior beliefs and the high educational level is the premium job and a wage equal to the required productivity in that job, $a_p(e_2) = \{\text{premium job}, \overline{\pi}\}$.

[10] With this interpretation, $\overline{\pi}$ includes the productivity of the worker in the premium job plus the positive synergy that the worker would generate if she held the premium job in the company with the second highest positive synergy. As workers are not usually ex-ante aware of the positive synergy they may generate in the companies where they end up working, we have assumed that companies are privately informed on this synergy, but the results are the same if we assume perfect information of that synergy. The presence of the positive synergy in the premium job guarantees that companies prefer the premium to the inferior job in the model.

Given those optimal responses, $a(t_1, e_1) \leq a_p(e_1) \leq a(t_2, e_1)$, $a(t_1, e_2) \leq a_p(e_2) \leq a(t_2, e_2)$ and $a_p(e_2) = a(t_2, e_2)$ as required in the general model.

In this simplified model, it is straightforward to prove that assumptions C1 and C2 hold because $\overline{\pi} > \pi_l(t_2, e_1) + c(t, e_2) - c(t, e_1)$. In fact, the high wage received in the premium job implies that the worker will be better off with the high level of education in the premium job than with the low level of education in the inferior job and the best possible wage in that job. Therefore, the first part of assumption C1 is satisfied. As the premium job is optimally offered by the company only if the worker has a high ability and a high level of education, the low-ability worker has strong incentives to imitate the high-ability worker and the second part of assumption C1 also holds. Lastly, as formal education only increases the worker′s productivity when she occupies the premium job, the worker prefers the high level of education as long as it allows her to access the premium job. Similarly, the worker prefers the low level of education as long as it prevents the company from treating her as the low-ability worker in the inferior job. For those reasons, assumption C2 is satisfied.

To finish off this section, we obtain our last result, which is a direct application of the results obtained in the previous section.

Proposition 2 *There are only two pooling PBE in pure strategies that survive divinity: One in which all workers' types choose e_1 and another in which they choose e_2. In addition to this, there are PBE in mixed strategies in which both worker's types choose e_2 with positive probability and they survive divinity. However, only the pooling PBE in pure strategies in which the worker chooses e_2 and the PBE in mixed strategies are NPE.*

For the sake of space, we omit the proof of this result because it is a direct application of the results obtained in the general model.[11] *In our model, the high-ability worker has strong incentives to choose the high level of education in order to access the premium job and receive a high wage. As the company's best response to the high level of education under the prior beliefs is offering the premium job and the high wage, the low-ability worker has also strong incentives to choose the high level of education. As a result, the pooling equilibrium in which the worker invests in the high level of education is the most plausible result, but standard refinements are not able to discard other equilibria in a universal or non-selective educational system because the single-crossing condition is not satisfied. For this reason, we apply the refinement introduced by Farrell* (1993) *that selects the most plausible equilibrium of this model. As shown in proposition* 1, *if we introduce a single-crossing condition by assuming that the cost of education decreases with the worker's productivity,*

[11] The proof of proposition 2 and a formal proof of the fact that assumptions C1 and C2 hold in this model are available from the author upon request. The online appendix includes an extension of the model with more than two possible types and more than two possible levels of education. As a numerical application of the model, let us assume that $\pi_l(t_1, e_1) = 1$, $\pi_l(t_1, e_2) = 4$, $\pi_l(t_2, e_1) = 8$, $\pi_l(t_2, e_2) = 10$, $\overline{\pi} = 12$, $c(t_1, e_1) = c(t_2, e_1) = 0$, $c(t_1, e_2) = c(t_2, e_2) = 1$, $p = 0.5$. With this example, we obtain the result shown in proposition 2.

then divinity selects the same overeducation pooling equilibrium, which is the most plausible equilibrium in this model. However, we have shown that the equilibrium selection proposed by Farrell (1993) is the same with and without this condition in our class of signaling games, which suggests that the neologism proof equilibrium is more robust to the specification of the single-crossing condition than divinity in this class of games.

7 Conclusions

In this article, we have characterized a class of signaling games in which an oversignaling equilibrium is the most plausible outcome. In that equilibrium, low sender′s types have incentives to imitate the highest type by sending the highest possible message. As expected, this kind of behavior will appear when the highest message induces the receiver to choose the sender′s favorite action and the consequences of being treated as a weak type are sufficiently adverse.

In our class of signaling games, a multiplicity of pooling equilibria arises, but the equilibrium in which the sender chooses the highest message is more plausible than others. In those games, there is a certain alignment of the sender′s and the receiver′s preferences. Specifically, when the receiver responds to messages with his prior beliefs, all sender′s types prefer a particular message. Furthermore, when the receiver responds to that message against his prior information, he chooses the action most preferred by all types of sender. As a result, all sender′s types will have strong incentives to choose that particular message in order to keep the receiver uninformed. Despite the plausibility of this result in our class of games, when the single-crossing condition is not satisfied, divinity and other standard refinements cannot eliminate implausible outcomes, but the refinement introduced by Farrell (1993) is able to select that plausible equilibrium. Thus, this paper has identified a class of signaling games in which the notion of equilibrium introduced by Farrell (1993) is more useful than other refinements.

In the literature on refinements, we have shown that divinity is stronger than other criteria in our class of signaling games because it selects a subset of the equilibria selected by other refinements. However, the undefeated equilibrium suggested by Mailath et al. (1993) cannot be compared with divinity this way because, in that refinement, the receiver′s beliefs after an out-of-equilibrium message must be equal to the posterior beliefs prescribed by an alternative equilibrium in which that message is in the equilibrium path. In our class of games, there are multiple pooling equilibria and each of the out-of-equilibrium messages is an equilibrium message in a different PBE in pure strategies and in some equilibrium in mixed strategies. In this context, all the equilibria are undefeated because the out-of-equilibrium beliefs are consistent with those prescribed by Bayes′ rule in a PBE in mixed strategies and as a result, this refinement does not help to select the oversignaling equilibrium.

As an application of our general class of signaling games, we have analyzed an extension of Spence′s model in which the educational investment increases the worker′s productivity in the labor market, and the highest level of education allows the worker to access a premium job with training opportunities. As expected, when the cost of education does not depend on the worker′s ability and the potential wage received in the premium job with the highest level of education is sufficiently high, there are multiple pooling equilibria, but the outcome in which all worker′s types choose the highest level of education is the only neologism-proof equilibrium in pure strategies.[12]

Appendix

Proof of Lemma 1 Let us consider one of those potential pooling equilibria in which all types send the same message, which can be any $m \in M$, R chooses his best response to that message, $a_p(m)$, and the on-path posterior beliefs equal the prior beliefs. Assumption C2 implies that no type would deviate towards a different message m' if R responded to the deviation as if the sender′s type is the lowest. As the out-of-equilibrium beliefs are unrestricted in a PBE, R′s beliefs might be concentrated on t_1 after observing m'.

Additionally, by construction, R′s choice is individually rational because $a_p(m)$ is a best response to S′s message when the posterior and prior beliefs are equal. Finally, as it is a pooling equilibrium, according to Bayes′ rule, the posterior beliefs in equilibrium have to be equal to the prior beliefs. Therefore, those beliefs are consistent.

So far, we have proven that there are N pooling equilibria in signaling games satisfying C2. Now, we prove that there is no other type of equilibria in pure strategies by contradiction. Imagine there were an equilibrium in which t_n sends $m \in M$, but this message is different from $m' \in M$, which is the message sent by $t_l < t_n$, were $t_l \in T$. Let I be the set of types which send m and J the set which sends m'. I and J may consist of a single type or more than one type. In any equilibrium in pure strategies, any set of types must send only one message with probability one and for this reason, $I \cap J = \emptyset$.

In the equilibrium we are considering, $r_m \in BR_{PB}(I, m)$ and $r_{m'} \in BR_{PB}\left(J, m'\right)$ denote R′s best responses to m and m', respectively. Then, part ii of C1 implies that there exists some $t \in J$ such that:

$$\overline{u}\left(t, m', r_{m'}\right) < \overline{u}\left(t, m, r_m\right) \tag{9}$$

[12] The online appendix shows an application of our class of signaling games to a model of advertising with false information. Additionally, that online appendix contains a generalization of our educational model with more than two worker′s types and more than two levels of education, and with a different version of the single-crossing condition. Besides that, that appendix includes a different version of Spence′s model without promotion opportunities. As the employer′s set of actions is a continuous interval in that application, there is only one assumption of our general class of signaling games that cannot be satisfied: $a_p\left(m_N\right) = a\left(t_n, m_N\right)$. In this application, we replace this assumption with an equivalent version and prove that the set of equilibria obtained and the type of equilibria selected by Farrell′s criterion are the same as those obtained in our general model. Finally, the online appendix contains an application of our general class of signaling games to the interaction between a mother and her child, which is an example of a cheap-talk game.

Thus, there would be a type belonging to J who prefers the message sent by I. Consequently, there is no PBE in which one type sends a different message from that sent by t_n. QED.

Proof of Lemma 2 We prove this lemma by contradiction. Imagine that there were a PBE which satisfies the three conditions of lemma 2. In this equilibrium, the utility obtained by a sender's type, $\mathrm{t} \in \mathrm{J}$, will be $\overline{\mathrm{u}}\left(\mathrm{t}, \mathrm{m}', \mathrm{r}_{\mathrm{m}'}\right)$, where $\mathrm{r}_{\mathrm{m}'} \in \mathrm{BR}\left(\mathrm{J}, \mathrm{m}'\right)$ is the mixed best response chosen by R after observing m′. It is easy to see that the following sequence of inequalities holds:

$$\begin{aligned} &\overline{u}\left(t, m', r_{m'}\right) \leq \overline{u}\left(t, m', r_{m'}^{PB}\right) < \overline{u}\left(t, m, r_{m}^{PB}\right) \leq \overline{u}\left(t, m, r_{m}\right) \forall r_{m'}^{PB} \\ &\in BR_{PB}\left(J, m'\right) \forall r_{m}^{PB} \in BR_{PB}(I, m) \end{aligned} \tag{10}$$

where $\overline{u}\left(t, m', r_{m'}^{PB}\right)$ is the utility obtained by a t-type of sender belonging to J when the receiver's posterior beliefs are induced by the prior probabilities concentrated on J, and $r_m \in BR(I, m)$ is the mixed best response to m chosen by the receiver in equilibrium. The first inequality in (10) is satisfied because $\mathrm{P_J(t)}$ first-order stochastically dominates $\mathrm{M_J}\left(\mathrm{t|m'}\right)$, $a\left(t, m'\right)$ is non-decreasing with t and $u(\cdot)$ increases with a. The second inequality in (10) follows from part ii of C1 because $\mathrm{t}_n \in \mathrm{I}$, $\mathrm{I} \cap \mathrm{J} = \emptyset$ and $\mathrm{m}' \neq \mathrm{m}$, which guarantees that there is at least a sender's type, $t \in J$, such that that inequality is satisfied. Finally, as $\mathrm{M_I(t|m)}$ first-order stochastically dominates $\mathrm{P_I(t)}$, $a(t, m)$ is non-decreasing with t and $\mathrm{u}(\cdot)$ increases with a, the last inequality in (10) holds.

Therefore, we conclude that $\overline{u}\left(t, m', r_{m'}\right) < \overline{u}\left(t, m, r_m\right)$, which implies that there is a type of sender belonging to J who would strictly prefer m to m'. This statement contradicts the rationality condition of this PBE. Then, our result has been proven. QED.

Proof of Theorem 1 This theorem follows immediately from Lemmas 1 and 2. QED.

Proof of Proposition 1 We consider pure and mixed strategy equilibria separately.

Pure Strategy Equilibria. Let us explain how we can discard a pooling equilibrium in pure strategies in which all types of sender send $m < m_N$ by using divinity. Under assumptions C1 and C2, for each $t \in T$, we obtain the following sequence of inequalities:

$$u\left(t, m_N, a(t_1, m_N)\right) \leq \overline{u}^*(t) = u\left(t, m, a_p(m)\right) \leq u\left(t, m, a(t_n, m)\right) < u\left(t, m_N, a(t_n, m_N)\right) \tag{11}$$

where the first inequality follows from assumption C2. The next inequality is due to the fact that $u(\cdot)$ increases with a and $a(t, m)$ is non-decreasing with t. The third inequality follows from part i in assumption C1. Then, the inequalities in (11) guarantee that for each $t \in T$, there exists some mixed best response to m_N: $r_0 \in P^0\left(t|m_N\right)$, and some mixed best response to m_N:$r_+ \in P\left(t|m_N\right)$, such that $\overline{u}^*(t) = u\left(t, m, a_p(m)\right) = \overline{u}\left(t, m_N, r_0\right)$ and $\overline{u}^*(t) = u\left(t, m, a_p(m)\right) < \overline{u}\left(t, m_N, r_+\right)$. Then, assumption C3 implies that for each $t \in T \setminus \{t_n\}$:

$$\overline{u}^*\left(t'\right) = u\left(t', m, a_p(m)\right) < \overline{u}\left(t', m_N, r_0\right) \forall t' > t \tag{12}$$

$$\overline{u}^*\left(t'\right) = u\left(t', m, a_p(m)\right) < \overline{u}\left(t', m_N, r_+\right) \forall t' > t \tag{13}$$

These inequalities allow us to conclude that $r_0, r_+ \in P\left(t'|m_N\right) \forall t' > t$ and consequently, $P\left(t|m_N\right) \cup P^0\left(t|m_N\right) \subset P\left(t'|m_N\right) \forall t' > t$. Thus, divinity requires that R attach positive posterior probability only to t_n after observing m_N. As a result, R's response to m_N must be $a(t_n, m_N)$ and we obtain that $\overline{u}^*(t) = u\left(t, m, a_p(m)\right) < u\left(t, m_N, a(t_n, m_N)\right)$ as shown by the last part of the sequence described in (11). This inequality implies that all the sender′s types will prefer to deviate to m_N and for this reason, no pooling equilibrium in pure strategies in which all types of sender choose $m < m_N$ survives divinity.

At this point, we show that the pooling equilibrium with m_N cannot be ruled out. We obtain:

$$\begin{aligned} \overline{u}^*(t) = u\left(t, m_N, a_p(m_N)\right) = u\left(t, m_N, a(t_n, m_N)\right) > u\left(t, m, a(t_n, m)\right) \\ \geq \overline{u}(t, m, r) \forall t \in T, \forall r \in BR(T, m) \text{ and } \forall m < m_N \end{aligned} \tag{14}$$

The first equality in (14) is the utility of the *t-type* of sender in this equilibrium. The second equality is due to the fact that $a_p\left(m_N\right) = a(t_n, m_N)$. The next inequality follows from part i of assumption C1 and the last inequality follows because the sender′s utility increases with a and $a(t, m)$ is non-decreasing with t. From (14), we deduce that no sender′s type can make a profit from a deviation to a lower message whatever the receiver′s response to that deviation. Thus, we cannot discard this equilibrium with divinity.

Now, we prove that no pooling equilibrium in pure strategies in which all types send a message m lower than m_N is a NPE. In this equilibrium, m_N is sent with probability zero, and when R's beliefs after that message are equal to the prior probabilities, R responds to m_N with $a_p\left(m_N\right) = a\left(t_n, m_N\right)$, and as shown by the last part of the sequence of inequalities (11), $u\left(t, m, a_p(m)\right) \leq u\left(t, m, a\left(t_n, m\right)\right) < u\left(t, m_N, a\left(t_n, m_N\right)\right)$. Hence, all types would deviate towards the highest possible message. Therefore, m_N is a credible neologism because we found a set of types $K = T$ such that,

$$\overline{u}^*(t) < u\left(t, m_N, a\left(t_n, m_N\right)\right) \forall t \in T \tag{15}$$

where $a\left(t_n, m_N\right) = BR_{PB}\left(T, m_N\right)$, which is unique in our model. As $T \setminus T = \emptyset$, inequality (2) of the definition of a NPE is trivially satisfied. Therefore, no pooling equilibrium in which all types send $m < m_N$ is a NPE. Furthermore, the sequence of inequalities in (14) also proves that the PBE in pure strategies in which all types of sender choose m_N is a NPE because no type can make a profit from a deviation. It is worth pointing out that we have not used assumption C3 in order to prove that the only NPE in pure strategies is the PBE in which all sender′s types choose m_N.

Mixed Strategy Equilibria. Here, we analyze the PBE in mixed strategies, but first of all, we prove that there cannot be an equilibrium in which none of the messages chosen by t_n with positive probability are chosen by another type. Assumption C1 guarantees that this equilibrium cannot exist. In fact, if all the messages chosen by t_n with positive probability were not chosen by any other sender′s type in a potential equilibrium, there would be a message, m, which is only sent by t_n, and a lower set of types, $J \subset T$, would be sending a different message, m', in that equilibrium. Consequently, the receiver would perfectly identify t_n after m. Let $r_{m'}$ be the receiver′s mixed best response to m' in that equilibrium. Then, it is easy to see that the following sequence of inequalities is satisfied for each $t \in J$:

$$\overline{u}\left(t,m',r_{m'}\right) \leq u\left(t,m',a\left(t_M,m'\right)\right) < u\left(t,m,a\left(t_n,m\right)\right) \tag{16}$$

where $t_M = \max_t J$. The first inequality in (16) follows because $a(t,m)$ is non-decreasing with t and $u(\cdot)$ increases with a, whereas the last inequality follows from part ii of assumption C1, where $I = \left\{t_n\right\}$, $J = \left\{t_M\right\}$ and m is one of the messages chosen by t_n in the equilibrium in question. Inequality (16) shows that types in J would strictly prefer m to m', which would contradict the rationality condition of the equilibrium in question. Therefore, we have proven that some of the messages chosen by the highest type must be chosen by at least some other type in equilibrium.

In the next step of our proof, we show that those PBE in mixed strategies in which no type sends m_N with positive probability do not survive divinity. Let us consider one of those equilibria in which a message, $m < m_N$, is sent by t_n and other types with positive probability. As shown above, such a message must exist. Then, $I \subseteq T$ will denote the set of types who send m with positive probability, including t_n. In this equilibrium, the utility received by a t-type belonging to I with message m will be represented by $\overline{u}^*\left(t,m,r_m\right)$, where $r_m \in BR(I,m)$. As no type has chosen m_N with positive probability in this equilibrium, for each $t \in I$, there must exist at least one $r_- \in BR\left(T,m_N\right)$ such that $\overline{u}^*\left(t,m,r_m\right) \geq \overline{u}\left(t,m_N,r_-\right)$. Otherwise, t would strictly prefer m_N to m for any mixed best response to m_N and would not have chosen m with positive probability. Next, it is easy to see that the following sequence of inequalities is satisfied for each $t \in I$:

$$\overline{u}\left(t,m_N,r_-\right) \leq \overline{u}^*\left(t,m,r_m\right) \leq u\left(t,m,a\left(t_n,m\right)\right) < u\left(t,m_N,a\left(t_n,m_N\right)\right) \tag{17}$$

The first inequality follows from the above analysis, the second holds because $a(t,m)$ is non-decreasing with t and $u(\cdot)$ increases with a and the last follows from part i of C1. The inequalities in (17) guarantee that for each $t \in I$, there exists some mixed best responses to m_N, $r_0 \in P^0\left(t|m_N\right)$ and $r_+ \in P\left(t|m_N\right)$, such that $\overline{u}^*\left(t,m,r_m\right) = \overline{u}\left(t,m_N,r_0\right)$ and $\overline{u}^*\left(t,m,r_m\right) < \overline{u}\left(t,m_N,r_+\right)$. Hence, assumption C3 implies that for each $t \in I \setminus \left\{t_n\right\}$:

$$\overline{u}^*\left(t',m,r_m\right) < \overline{u}\left(t',m_N,r_0\right)\forall t' > t \tag{18}$$

$$\overline{u}^*\left(t',m,r_m\right) < \overline{u}\left(t',m_N,r_+\right)\forall t' > t \tag{19}$$

These inequalities allow us to conclude that $r_0, r_+ \in P(t'|m_N)\forall t' > t$ and consequently, $P(t|m_N) \cup P^0(t|m_N) \subset P(t'|m_N)\forall t' > t$ Thus, divinity requires that the receiver attach positive posterior probability only to $t_n \in I$ after observing m_N. As a result, R's response to m_N must be $a(t_n, m_N)$ and we obtain that $\overline{u}^*(t, m, r_m) < u(t, m_N, a(t_n, m_N))$ as shown by the last part of the sequence described in (17). This inequality implies that all sender's types in I will prefer to deviate to m_N and for this reason, an equilibrium in which no type chooses m_N with positive probability is not divine.

In the last step of our proof, we show that an equilibrium in mixed strategies in which no sender's type sends m_N with positive probability cannot be a NPE. In that equilibrium in mixed strategies, m is some message sent by some types with positive probability and r_m is the best response to that message. Imagine that R's beliefs are equal to the prior probabilities after m_N. In this case, R's response to m_N should be $a_p(m_N) = a(t_n, m_N)$ and as shown by the sequence of inequalities included in (14), $u(t, m_N, a(t_n, m_N)) > u(t, m, a(t_n, m)) \geq \overline{u}(t, m, r)\forall r \in BR(T, m), \forall t \in T$. Thus, m_N would be a credible neologism and we conclude that the equilibrium in question is not a NPE.

Summing up, we conclude that D1 and Farrell's refinement select the same equilibria in our class of games when assumptions C1-C3 are satisfied, but we did not need assumption C3 in order to select those equilibria with Farrell's criterion. QED.

Proof of Theorem 2 The proof of this theorem is shown in the above proof of proposition 1 where we selected the NPE without assumption C3. QED.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00712-024-00882-x.

Acknowledgement I am very grateful to all participants in the SING annual conference held in Finland and in the seminars held at the departments of economics of the Universities of Crete, Lancaster, Nottingham, Surrey and Warwick for their useful suggestions. This article has also been benefited from fruitful discussions with professor Alex Possajennikov and Martin Jensen at the University of Nottingham. Furthermore, comments and suggestions made by Facundo Albornoz, María Montero, Silvia Sonderegger and Spiros Bougheas have helped improve the quality of the paper. Finally, I thank several anonymous referees for their helpful comments. All errors and omissions are my own. Part of this research was supported by the University of Malaga (Plan Propio), which partially financed my research visits to the University College London, the University of Warwick, the University of Nottingham and Lancaster University, where part of my research was carried out.

Funding Funding for open access publishing: Universidad Málaga/CBUA.

Declarations

Conflict of interest I would like to confirm that there are no ethical issues and conflict of interests related to this study.

References

Cho I-K, Kreps DM (1987) Signaling games and stable equilibria. Q J Econ 102(2):179–202. https://doi.org/10.2307/1885060
Cho I-K, Sobel J (1990) Strategic stability and uniqueness in signalling games. J Econ Theory 50:381–413. https://doi.org/10.1016/0022-0531(90)90009-9
Clark D, Fudenberg D (2021) Justified communication equilibrium. Am Econ Rev 111(9):3004–3034. https://doi.org/10.1257/aer.20201692
Dominiak A, Lee D (2023) Testing rational hypothesis in signaling games. Eur Econ Rev 160:1–27. https://doi.org/10.1016/j.euroecorev.2023.104610
Esó P, Schummer J (2009) Credible deviations from signalling equilibria. Int J Game Theory 38:411–430. https://doi.org/10.1007/s00182-009-0161-x
Farrell J (1993) Meaning and credibility in cheap-talk games. Games Econ Behav 5(4):514–531. https://doi.org/10.1006/game.1993.1029
Flower T (2011) Fork-tailed drongos use deceptive mimicked alarm calls to steal food. Proc R Soc B 278:1548–1555. https://doi.org/10.1098/rspb.2010.1932
Fudenberg D, He K (2017) Type-compatible equilibria in signalling games. https://pdfs.semanticscholar.org/6fdb/83b03341e89dfac9da93eb025c2d9e3bfb20.pdf. Accessed 5 September 2018.
Fudenberg D, He K (2018) Learning and type compatibility in signaling games. Econometrica 86(4):1215–1255. https://doi.org/10.3982/ECTA15085
Fudenberg D, He K (2020) Payoff information and learning in signaling games. Games Econ Behav 120:96–120. https://doi.org/10.1016/j.geb.2019.11.011
Gibbons R, Waldman M (1999) A theory of wage and promotion dynamics inside firms. Q J Econ 114(4):1321–1358. https://doi.org/10.1162/003355399556287
Grossman SJ, Perry M (1986) Perfect Sequential Equilibrium. J Econ Theory 39(1):97–119. https://doi.org/10.1016/0022-0531(86)90022-0
Johnson JP, Myatt DP (2006) On the simple economics of advertising, marketing and product design. Am Econ Rev 96(3):756–784. https://doi.org/10.1257/aer.96.3.756
Kihlstrom RE, Riordan MH (1984) Advertising as a Signal. J Pol Econ 92(3):427–450. https://doi.org/10.1086/261235
Kohlberg E, Mertens JF (1986) On the strategic stability of equilibria. Econometrica 54(5):1003–1037. https://doi.org/10.2307/1912320
Liu S, Pei H (2020) Monotone equilibria in signalling games. Eur Econ Rev 124:1–16. https://doi.org/10.1016/j.euroecorev.2020.103408
Mailath GJ, Okuno-Fujiwara M, Postlewaite A (1993) Belief-based refinements in signaling games. J Econ Theory 60(2):241–276. https://doi.org/10.1006/jeth.1993.1043
McLennan A (1985) Justifiable beliefs in sequential equilibrium. Econometrica 53:889–904. https://doi.org/10.2307/1912659
Milgrom P, Roberts J (1982) Limit pricing and entry under incomplete information. Econometrica 50(2):443–459. https://doi.org/10.2307/1912637
Milgrom P, Roberts J (1986) Price and advertising signals of product quality. J Pol Econ 94(4):796–821
Pei H (2020) Reputation effects under interdependent values. Econometrica 88(5):2175–2202. https://doi.org/10.3982/ECTA16584
Riley JG (1979) Informational equilibrium. Econometrica 47(2):331–359. https://doi.org/10.2307/1914187

Sobel J, Stole L, Zapater I (1990) Fixed-equilibrium rationalizability in signaling games. J Econ Theory 52(2):304–331. https://doi.org/10.1016/0022-0531(90)90035-I
Spence AM (1973) Job market signaling. Q J Econ 87(3):355–374. https://doi.org/10.2307/1882010