Overeagerness*

Philipp Sadowski†

July 2016

Abstract

We capture the impression that high types may send lower signals than low types in order not to appear too desperate. We require a noisy one-dimensional signal, where a very low signal being transmitted forces types to execute their outside option. The central assumption is that low types are not only less productive when employed, but that they also face a worse outside option. High types then exploit low types' eagerness not to end up with their bad outside option by running a larger risk of transmitting a very low signal.

1. Introduction

With his classical signaling model, Spence (1973) illustrates circumstances under which rational agents should engage in a wasteful activity in order to distinguish themselves from less competitive individuals. In his terminology, "high types" send a costly but non-productive "signal" that "low types" cannot justify sending. Though hard to verify empirically,[1] signaling has been considered one reason for phenomena like people acquiring educational degrees, job candidates wearing a nice outfit for their interview, or companies advertising their products.

Surprisingly, we sometimes observe ourselves reacting negatively to such signaling behavior. For example, heavy advertisement of a product, say in the form of many friendly phone calls, may not improve our view of its quality, and in a rare field experiment Bertrand et al. (2010) find that adding an additional promotional feature to a loan offer my reduce the loan take-up rate (see Section 3 for details).

If this type of reaction is justified, we should find that sometimes the less competitive send stronger signals than the most competitive. Indeed, some luxury brands, like good

*I thank Roland Benabou, Avinash Dixit, Tymon Tatur, Huseyin Yildrim, two anonymous referees, and the editor for useful comments.

†Department of Economics, Duke University. E-mail: p.sadowski@duke.edu

[1] See Weiss (1995) for a convincing attempt in the context of wages.

wineries, hardly advertise at all. In the context of job-market signaling, Hvide (2003) cites casual evidence of the most capable of students leaving or skipping college and going straight into business in areas where education is not a formal requirement for entry. For example, Stanford University is known to have lost a substantial amount of students to the high-tech sector. Also, Orzach and Tauman (1996) note that in the 1996 Forbes 400 list containing the richest 400 people in the United States, very many do not have any academic degree.[2] Furthermore Feltovich et al. (2002) point out, that in the US talented students (as measured by aptitude tests such as the SAT) tend to underachieve in terms of school grades. These examples are only suggestive, but the notion that eagerness is often interpreted as a bad sign and that lack of eagerness can be a positive indicator is also a conventional wisdom: Modesty is attractive when, in essence, it describes individuals who seem less concerned about impressing others, whereas boasting may well have a negative impact on the perception of an individual's capabilities. The proverb "Barking dogs don't bite" captures the idea well.

This paper provides a signaling model that is consistent with the impression that eagerness comes from desperation or a bad outside option, which is a negative sign, whereas lack of eagerness suggests that the sender must have a better outside option, which is a positive sign.

We make three main assumptions. First, we assume that high types are distinguished from low types not only by being more productive when employed (and possibly by finding it cheaper to send a high signal), but also by a higher opportunity cost of being employed. For example, they may also be more productive when self employed, giving them a better outside option.[3]

Second, we assume that individuals have only a noisy signal available to them. Higher signal-sending effort results in a higher signal, in the sense that it reduces the risk of a low signal manifestation.

Third, very low signal manifestations force individuals to execute their outside option. The content of this assumption depends on the application:

- The signal amplitude might have to exceed a physical threshold in order to be detectable. For instance, an advertisement with a subdued color palette may be more likely to go unnoticed than one with bold colors.

[2] In both examples the decision to leave school may have occurred after an employment option presented itself. However, if a good degree has signaling value even for successive employers, one could argue that such young professionals made the career decision to be better off without such a signal.

[3] That not only the absolute advantage in terms of productivity when employed is relevant, but also how this relates to productivity in another activity, here self-employment, is akin to the importance of comparative advantage (as opposed to absolute advantage) in theories of trade. That said, the mechanics of a market equilibrium with multiple goods are very different from those of the present signaling model.

- There could be three types in the population, unskilled types, and skilled types of low or high ability, where the large majority is unskilled, does not have any productive value for the company, and is unable to send a high signal.[4] For instance, some essential skill may have to have been acquired over time and independent of ability, and an available exam may be such that unskilled individuals can not answer any question correctly, while the number of questions skilled types expect to answer correctly depends on effort, as formalized in Example 1.

The two scenarios are equivalent. We formulate our model according to the second, but Section 3 discusses evidence that better fits the first.

The three features of our model make skilled individuals of low type very eager not to appear unskilled, because they would face their unattractive outside option. They tend to send a high signal to reduce this risk. Skilled individuals of high type, on the other hand, are less concerned about being seen as unskilled, because their outside option is more attractive. If high types do accept employment, however, they would like to be distinguishable from low types so they can be rewarded according to their productivity. Hence, they may choose to send a lower signal, allowing some distinction between types, though at the cost of increasing their risk to be perceived as unskilled. This is exactly the kind of behavior we hope to illustrate in a model of unproductive signaling.

Spence's assumption of a (type dependent) cost of sending a signal can easily be incorporated. In standard signaling contexts, this assumption is central, as it prevents low types from perfectly imitating high types. It drives none of the features of our model. On the contrary, if signaling is cheaper for high types, then the classical signaling effect compensates overeagerness. Furthermore, in our setup it is not clear which type should incur the higher loss in utility from sending a given signal. For example, high types may need to spend fewer units of time on acquiring a certain educational degree, but they might also value each unit of time higher due to their better outside option.[5] In particular, our model applies to situations where cheap talk is possible for both types. Accounting for type dependent costs of

[4] It is essential that unskilled types only have access to "lower" signal distributions than skilled types. The stronger assumptions that the support of their signal distribution is bounded above and that they constitute the vast majority of potential applicants simplify the analysis, as in that case firms will only employ applicants who can demonstrate skill. Unskilled individuals then have no incentive to behave strategically, and the analysis can focus on the behavior of the two skilled types.

[5] Similarly, consider the example of advertising: It may be easier to highlight the positive features of a high quality product. For example consumer tests can be cited, etc. In that case, the same level of advertising would cost less money for high quality producers. At the same time, a high quality producer might make better use of every dollar he does not invest in a public advertising campaign, say by investing in the relations with his existing customer base, the same base that constitutes his better outside option to a successful campaign. Therefore, taking opportunity costs into account, advertising might actually be cheaper for a low type.

sending the signal allows us to predict when overeagerness should occur, which should, in principle, be testable.

There are other models that feature signals that are not always monotonic in quality. For example, Teoh and Hwang (1991) show that a firm of high type may withhold good news from investors, whereas a firm of low type, with a bleaker outlook on the future, would disclose the same news. Orzach et al. (2001) explain how "Modest advertising signals strength," if firms have multiple periods to sell a product and use price and advertisement expenditure as a two dimensional signal. Similarly, Clements (2004) describes how the quality of a product's packaging can be non-monotonic in the quality of the product, if price is used as a signal, as well. Araujo et al. (2004) address the observation that wages are non-monotonic in the GED of high school dropouts. Feltovich et al. (2002) show that signaling can be non-monotonic if there are three types and another exogenous dimension of the signal is available. When non-monotonic signals occur in the context of such multidimensional signals they are also referred to as counter-signaling. Daley and Green (2014) also consider a two dimensional signal where one dimension is exogenous and show that a high initial reputation may lead to lower signals. Non-monotonic signals also arise in other contexts: Benabou and Tirole (2004) construct a reversal of high and low signals in the context of pro-social behavior by considering type dependent preferences: if signaling brings a direct monetary reward, then the most altruistic individuals may find it more costly to signal altruism than individuals who are greedy. Chung and Esö (2013) have a two period signaling model with non-monotonic equilibria, where signals are ordered according to being more or less informative, rather than being good or bad. In contrast to these models, our explanation of non-monotonic signals applies also in situations that involve only one period and a one-dimensional signal, as for example the field experiment by Bertrand et al. (2010) referred to above and discussed in Section 3.[6]

The basic version of our model allows cheap-talk, features a noisy signal, and relies on the presence of unskilled types, which are essentially non-strategic senders. The cheap-talk literature has previously considered noisy communication channels, for example Blume et al. (2007), as well as non-strategic senders, for instance the honest senders in Chen (2011).

Section 2 of this paper introduces our model of overeagerness and contains the general result that low types will engage more heavily in signaling than high types in any equilibrium where high types do not send the highest possible signal. Section 3 discusses an application to advertising. Section 4 introduces type dependent cost of effort. Section 5 concludes.

[6] Due to the noise there can be partially separating equilibria, allowing us to meet the incentive constraint with a one-dimensional signal.

2. A Model of Overeagerness

We consider a sender-receiver game. Senders will also be called applicants, receivers firms. There are three types of applicants: High types, low types, and unskilled types. High types have productivity $\theta_H \in \mathbb{R}_+^*$, low types have $\theta_L \in \mathbb{R}_+^*$ where $\theta_H > \theta_L > 0$, and unskilled types lack an essential skill, and hence have zero productivity when employed by the firm. High types have opportunity cost v_H for being employed by the firm, low types have opportunity cost v_L, where $v_H > v_L > 0$ and $\theta_L > v_L$. These parameters are common knowledge.

There is a continuum of potential applicants and common knowledge about the fraction r of the population that has the essential skill. The distribution of applicant types in the skilled fraction of the population is commonly known, too: The ratio of high types over low types is n. Let $\overline{\theta} := \dfrac{n\theta_H + \theta_L}{n+1}$ be the average productivity of all skilled applicants.

Skilled applicants can choose the effort $e \in [0,1]$ they put into sending a signal, where the signal value is a random draw from $S \subset \mathbb{R}$ according to the probability function $f_e(s)$, which has full support on S for all e. Either S is an interval and $f_e(s)$ is a density function, or S is finite and $f_e(s)$ is a mass function. Let $f_e(s)$ satisfy the Monotone Likelihood Ratio Property (MLRP): $\forall s_1 > s_2$ the ratio $\dfrac{f_e(s_1)}{f_e(s_2)}$ is increasing in e. Also let $f_e(s)$ be continuous in e for all $s \in S$. We assume for now that effort is free.[7]

Applicants without the required skill send signals from a distribution with probability (density or mass) function $\tilde{f}_u(s)$, which supports all $s \in S$ with $s \leq s^*$ for some $s^* \in S$. Let $\underline{s}$ and $\overline{s}$ be the smallest and largest elements of S, respectively. If S is an interval we require $\underline{s} < s^* < \overline{s}$. If S is finite, there must be $s \in S$ such that $s^* < s < \overline{s}$.

Example 1. *Applicants can generate a noisy, public and cost free test result. For simplicity, let the test consist of an infinite number of questions. Each can be answered right or wrong. After the test is completed, two questions are chosen at random for evaluation. So the test result $s \in S = \{0,1,2\}$ counts the correct answers to the two questions considered. For unskilled applicants, none of the questions can ever be answered correctly, that is, $s^* = 0$. Both types of skilled applicants can choose effort $e \in [0,1]$ to adjust the rate at which they answer each question correctly between a lower bound $\pi_* > 0$ and an upper bound $\pi^* < 1$ with $\pi^* > \pi_*$. Let $\pi : [0,1] \longrightarrow [\pi_*, \pi^*]$ be increasing such that $\pi(e)$ is that rate. Test signals are then distributed binomially for skilled applicants; given π, the probability not to get any question right ($s = 0$) is $p_0 = (1-\pi)^2$, the probability to get one of the questions right*

[7] Applicants are categorized into high and low types according to their opportunity costs, $v_H > v_L$, and it is important that two applicants of the same type face the same cost. The assumption that applicants of the same type are also equally productive is only for expositional convenience; one can think of θ_H and θ_L as the *average* productivity of types with high and low opportunity costs, respectively.

($s = 1$) is $p_1 = 2(1-\pi)\pi$ and the probability to get both questions right ($s = 2$) is $p_2 = \pi^2$.

For $i \in \{L, H\}$ the utility of a representative applicant is $u_i\left(w \,|\text{accept offer}\right) = w$ from accepting employment at a certain wage w or $u_i\left(w \,|\text{do not accept offer}\right) = v_i$ from executing their outside option.

On the firm's side there is perfect competition. Consequently, the wage offered contingent on signal s is the expected productivity of the largest subset of types that would accept such a wage.[8] The wage schedule offered by firms is $w : S \to [0, \theta_H]$.

We consider only pure strategies for the applicants: If indifferent between accepting and declining an offer, they are assumed to accept, and when indifferent between lower and higher effort, they are assumed to exert lower effort. A strategy is then an effort choice e, and a Nash equilibrium consists of effort choices (e_L, e_H), where the resulting wage for $s > s^*$ is

$$w\left(s\right) = \frac{nf_{e_H}\left(s\right)\theta_H + f_{e_L}\left(s\right)\theta_L}{nf_{e_H}\left(s\right) + f_{e_L}\left(s\right)}$$

if high types accept such an offer and $w\left(s\right) = \theta_L$ if they would decline it, and where applicants maximize their expected utility

$$U_i\left(w, e\right) = \int_S \max\left(w\left(s\right), v_i\right) f_e\left(s\right) ds$$

For $s \leq s^*$ wages are specified accordingly, taking into account the presence of unskilled applicants and whether or not the two skilled types would accept such an offer. The fraction $1 - r$ of unskilled applicants is assumed to be large in the sense that any justifiable offer would be unacceptable to low and high types. So $w\left(s\right) < v_L < v_H$ has to hold for $s \leq s^*$ in equilibrium.[9] Expected utility then becomes

$$U_i\left(w, e\right) = F_e\left(s^*\right) v_i + \int_{\{s \in S : s > s^*\}} \max\left(w\left(s\right), v_i\right) f_e\left(s\right) ds$$

where F_e denotes the cumulative distribution function corresponding to f_e.

[8] Due to the full support of the signal distribution, there is no out of equilibrium signal value. Due to zero profit firms can, then, be understood as a mechanism that rewards applicants.

[9] Specifically, it has to be true that high types want to deviate from accepting a wage justified by all types accepting, and low types want to deviate from accepting a wage justified by only skilled low and unskilled types accepting. Suppose, for simplicity, that unskilled types face opportunity cost $v_u = 0$. Then

$$\frac{1-r}{r} > \max\left(\frac{nf_{e_H}\left(s\right)\left(\frac{\theta_H}{v_H} - 1\right) + f_{e_L}\left(s\right)\left(\frac{\theta_L}{v_H} - 1\right)}{\left(n+1\right)\tilde{f}_u\left(s\right)}, \frac{f_{e_L}\left(s\right)\left(\frac{\theta_L}{v_L} - 1\right)}{\left(n+1\right)\tilde{f}_u\left(s\right)}\right)$$

for all $e_H, e_L \in [0, 1]$ and $s \leq s^*$ is a sufficient constraint on r to guarantee this.

Proposition 1. *In any Nash equilibrium of the model,* $1 \geq e_L > e_H \geq 0$ *or* $e_L = e_H = 1$.

All proofs can be found in the appendix. The intuition for this result was described in the introduction: skilled individuals of low type are very eager not to appear unskilled, because they would face their unattractive outside option. So they will tend to send a high signal to reduce this risk. Skilled individuals of high type on the other hand, are less concerned about being seen as unskilled, because their outside option is more attractive. If high types do accept employment, however, they would like to be distinguishable from low types so they can be rewarded according to their productivity. Hence, they may choose to send a lower signal, allowing some distinction between types at the cost of increasing their risk to be perceived unskilled.[10]

The next proposition characterizes equilibrium wages for the partially separating equilibrium.

Proposition 2. *In every equilibrium with* $e_H < 1$ *either skilled high types accept for any signal manifestation* $s > s^*$ *and*

$$w(s) = \begin{cases} \leq v_L & s \leq s^* \\ \dfrac{n f_{e_H}(s)\,\theta_H + f_{e_L}(s)\,\theta_L}{n f_{e_H}(s) + f_{e_L}(s)} & s^* < s \end{cases}$$

or there is $s^{**} \in S$ *with* $s^{**} \geq s^*$ *such that skilled high types accept only for* $s \in S$ *with* $s^* < s < s^{**}$ *and*

$$w(s) = \begin{cases} \leq v_L & s \leq s^* \\ \dfrac{n f_{e_H}(s)\,\theta_H + f_{e_L}(s)\,\theta_L}{n f_{e_H}(s) + f_{e_L}(s)} & s^* < s < s^{**} \\ \theta_L & s \geq s^{**} \end{cases}$$

Proposition 3. *An equilibrium always exists, and* $e_L = e_H = 1$ *is an equilibrium if and only if* $\bar{\theta} > v_H$.

Proposition 3 implies that for $\bar{\theta} \leq v_H$ an equilibrium with $e_L > e_H$ exists. We now reconsider the example from above to investigate the welfare implications of this equilibrium, and to claim that an equilibrium with $e_L > e_H$ may also exist for $\bar{\theta} > v_H$.

[10] Of course, low types would like to imitate high types in our model. Thus, if any separation occurs in equilibrium it has to be true that, taking everything but the wage upon employment into account, the expected cost a low type would incur when sending the noisy signal high types send has to exceed the cost of the noisy signal other low types send. Proposition 1 merely establishes that a positive correlation between type and outside option can reverse the role of high and low noisy signals, where we refer to the noisy signal with the lower probability of a low signal manifestation as the higher signal. When effort is costly, this could also be the signal bearing the higher direct cost.

Example 2 (Example 1 continued). *Suppose $\bar{\theta} \leq v_H$. Consider a situation with only a binary cost-free signal, $S = \{0, 1\}$, where unskilled types always send the low signal, $s = 0$, and the probability that skilled types send $s = 0$ is bounded below by $(1 - \pi_*)^2$. In the only equilibrium of this model skilled high types never accept employment and skilled low types work at $w = \theta_L$ with probability $1 - (1 - \pi_*)^2$. Now recall the model from Example 1, which features the same lower bound on the probability of a very low signal manifestation. In terms of welfare, the equilibrium with $e_L > e_H$, which exists by Proposition 1, is a strict Pareto improvement over the only equilibrium with a binary signal: high types sometimes find profitable employment, and with positive probability low types receive a higher wage than is justified by their own productivity.*

Example 3 (Example 1 continued). *We claim that, in Example 1, there exists $v^* > \bar{\theta}$ such that for $v_H \in \left(\bar{\theta}, v^*\right)$ there exists an equilibrium where skilled high types accept only w_1, $e_H = 1/2$, $w_2 = \theta_L$, and e_L and w_1 are jointly determined by zero profit for the firm and the low type best responding. By Proposition 1 then $e_L > e_H$. A proof is in the appendix.*

The analysis of signaling models a la Spence usually proceeds by separating the effects of type dependent costs of sending a signal, e, and the wage, w, on utility. If, given w, the marginal cost of improving the signal is always higher for low types, then payoffs satisfy the Spence-Mirrlees Single Crossing Property and $e_H \geq e_L$ must hold in equilibrium. In contrast, in our model the thought exercise of separating wage and cost is not useful: The marginal cost of changing a signal is the marginal probability of having to execute the outside option.[11] If offered a fixed wage, an individual would either accept employment or choose the outside option. In both cases, there would be no marginal cost of changing the signal. Separating cost from expected wage does not help, either: effort e and expected wage do not enter the utility function separately. Rather, e determines the distribution of signal manifestations $f_e(s)$. Thus, a change in the cost of a signal, which requires a change in $f_e(s)$, generally corresponds to a change in the expected wage.

3. An Application: Advertising

A bank is trying to acquire new customers by sending out a brochure about their current extremely competitive credit offer. The informative part of the brochure has a table of numbers explaining the offer. Due to the many offers of such kind, potential clients will only notice the offer, if the brochure is sufficiently appealing in a combination of aspects ($s > s^*$). Imagine one possibility to make the brochure appealing is to announce giving away mobile

[11] In addition there may be direct cost, as considered in Section 4.

phones to some readers. The bank does not know for sure, whether the brochure is already appealing enough to catch readers attention without the announcement.

There are two types of clients: High frequency borrowers have access to credit at terms which are only slightly worse than the credit on offer. They would not take up the credit on offer if they knew that the bank provided low quality service, θ_L. Low frequency borrowers, who have no comparable credits available to them, would accept the offer, even from a low quality bank. Potential new clients cannot observe the quality of the bank, θ_L or $\theta_H > \theta_L$, before deciding whether to accept the offer.

A high quality bank can thrive even without marketing effort to acquire new clients, because of word of mouth (outside option v_H). A low quality bank does not have word of mouth working for it ($v_H > v_L$). Its only option to thrive is to succeed in the marketing effort.

In this situation, it seems reasonable that a low quality bank would add the mobile-phone-give-away to its brochure to make sure it is as appealing as possible and receives attention (it expects to send a high signal s). Then, low frequency borrowers will most likely take the time to read the good offer and take it.

A high quality bank may then distinguish itself by not including a similar give-away on its brochure (it expects to send a lower signal). In that case, even though the brochure with the give-away engages their attention, high frequency borrowers may rightfully interpret such a brochure as a signal of a low quality bank and borrow money from their alternative source instead. Thus the high quality bank is running the risk of sending out an unappealing brochure that does not attract customers. However, if it manages to create an appealing brochure without a phone-give-away, it attracts low and high frequency borrowers (corresponding to a high wage in the setup of the model). The low quality bank tries to make sure it reaches the low frequency borrowers at the cost of not acquiring high frequency borrowers (a lower wage in the setup of the model).

The above scenario closely resembles a feature of the field experiment conducted by Bertrand et al. (2010) in South Africa, which was mentioned in Section 1. Without claiming that overeagerness explains their finding or that the description above does justice to the complexity of their experiment, it is worth noting the parallels: They do find a negative impact of such a phone-give-away on the take-up rate among high frequency borrowers, but not among low frequency borrowers. So, it does seem as if the give-away is interpreted as a signal of low quality among high frequency borrowers. They write that "…when we break up the sample into borrowing categories, we see that this effect [of the phone-give-away] is very large and statistically significant among the more frequent borrowers. For this group of customers, introducing this promotional feature […], in fact, reduces the likelihood of loan take-up. The nonnegative effect among the lower frequency borrowers may indicate that this negative choice effect of the promotional lottery may be offset, in this case, by an attention-getting-effect ..." This matches the behavior we suggest for the respective customers very well.

4. Type dependent cost of effort

To this point cheap talk is possible in our setup. We now incorporate type-dependent cost of effort for sending a certain signal. For example, preparation for an exam might be easier for one type than the other.

Signaling models typically assume the cost of sending a specific signal to be negatively correlated with productivity θ. Indeed, it seems reasonable to assume that high types would need less time to prepare a certain exam or, in general, send a particular signal. In our model, however, they also have a better outside option. Hence, it is plausible that the opportunity cost for spending a given amount of time would be higher for them than for low types. In our model both types draw their signal from the same distribution when exerting the same effort, so effort is measured as "output" in terms of signal send, not "input" in terms of time spent sending it. Hence, the opportunity cost of the time needed to exert a certain effort is the relevant cost. Therefore, it is unclear which type incurs the higher cost of effort.[12]

If the cost of effort is higher for low types we expect the classical signaling effect to compete with overeagerness. In principle, incorporating costs into the setup makes the model more applicable and testable. It allows one to predict whether or not to expect overeagerness.

To incorporate costs of effort, assume that there is a function $c(e), c : [0, 1] \to [0, \infty]$ with low type's cost $c_L(e) = c(e)$ and high type's cost $c_H(e) = ac(e)$ where $a \in (0, \infty)$. Let c be twice continuously differentiable, with $\frac{\partial c(e)}{\partial e} > 0$ and $\frac{\partial^2 c(e)}{\partial e^2} > 0$ for all $e \in (0, 1)$.

For parameter values where an equilibrium exists and $e_L > e_H$ in *all* equilibria, we shall say overeagerness dominates. For values where an equilibrium exists and $e_L < e_H$ in *all* equilibria, we say the classical signaling effect dominates. We say there is perfect pooling, if $e_L = e_H$ in equilibrium.

[12] See footnote 5 to the Introduction for an example in the context of advertising.

Proposition 4. *In the model specified above there is* $\bar{a} < 1$ *sufficiently large such that for* $a > \bar{a}$ *overeagerness dominates. Furthermore, if costs are significant enough in the sense that* $c'(0) > \underline{c}'$ *for an appropriate* $\underline{c}' > 0$*, then there is* $\underline{a} > 0$ *sufficiently small such that for* $a < \underline{a}$ *the classical signaling effect dominates.*
Perfect pooling can occur if and only if $a = a^* := \dfrac{\bar{\theta} - v_H}{\bar{\theta} - v_L} < 1$ *and* $\bar{\theta} > v_H$*.*

Proposition 4 immediately implies $\underline{a} \leq \bar{a}$ with strict inequality for $\bar{\theta} > v_H$. To gain intuition for the possible coexistence of distinct types of equilibria, consider the case where a perfect pooling equilibrium exists, $a = \dfrac{\bar{\theta} - v_H}{\bar{\theta} - v_L}$ and $\bar{\theta} > v_H$. In that equilibrium it is a strict best response for both skilled types to accept any wage offer for signal manifestations above s^*, and for each type the cost equals the expected benefit of employment.[13] Now note that, for $F_0(s^*)$ (which is the upper bound on the weight given to signal manifestations below s^*) large enough, and for $v_H > \theta_L$, there is clearly another equilibrium where $w(s) = \theta_L$ for all $s > s^*$, high types never accept a wage offer, $e_H = 0$, and low types choose $e_L > 0$, so that $e_L > e_H$.

5. Conclusion

The model presented here shows that high types may choose to send lower signals than low types in a framework very close to classical signaling models. Our predictions are driven by a correlation between individual outside option or opportunity cost and productivity when employed and by the assumption that signals are noisy. Firms correctly interpret very high signals as overeagerness of low types, who are desperate not to end up with their bad outside option. Only high types can afford the risk of a very low signal for the benefit of not being seen as overeager, because if they do end up with a very low signal they execute their own outside option, which is more attractive.

The predictions of our model, like those of Spence's original model, will be hard to test, mainly because the effect is difficult to isolate, because there are few completely unproductive signals etc. In principle, including type dependent cost of effort allows prediction of when overeagerness should occur, adding testability.

Consider the policy implications of the model in the context of workers applying for employment. From the perspective of a social planner, high types taking the risk of executing their outside option obviously impose a negative externality on low types, who are paid less

[13] Note that, by continuity, an equilibrium where both types accept all wage offers for $s > s^*$ survives when slightly reducing a and holding all else fixed. Since high types' cost of sending a signal is now proportionally lower, one can easily find examples of signaling technologies f_e, such that high types' best response must now be higher than low types' best response, so that $e_H > e_L$.

on average. If high types gain little from their behavior, this may be an argument for a society to prevent behavior as in the model, even if the signal is nonproductive. For example, if high types are willing to work for the wage justified by average productivity ($\bar{\theta} \geq v_H$), then formal degree requirements for a certain profession would improve welfare.

On the other hand, as we show in Example 2, rationally rewarding low signals may allow signal-specific wage offers to induce a proportion of capable individuals to work, who would decline employment at a uniform wage ($\bar{\theta} < v_H$). In this case giving up such formal requirements can lead to a *strict* Pareto improvement, as high types benefit and low types are paid more on average.

6. Appendix

Proof of Proposition 1.

Consider two cases:

i) Assume $e_H > e_L$. The wage schedule cannot be weakly increasing everywhere, as low types would deviate by playing $e_L = 1 \geq e_H$. So there are $s_1 > s_2$ with $w(s_1) < w(s_2)$. This holds if and only if the wage justified by both types accepting at signal s_1 is smaller than the one justified by both types accepting at s_2:

$$\frac{n f_{e_H}(s_1)\,\theta_H + f_{e_L}(s_1)\,\theta_L}{n f_{e_H}(s_1) + f_{e_L}(s_1)} < \frac{n f_{e_H}(s_2)\,\theta_H + f_{e_L}(s_2)\,\theta_L}{n f_{e_H}(s_2) + f_{e_L}(s_2)}$$

which is equivalent to

$$\frac{f_{e_H}(s_1)}{f_{e_L}(s_1)} < \frac{f_{e_H}(s_2)}{f_{e_L}(s_2)} \quad \Leftrightarrow \quad \frac{f_{e_H}(s_1)}{f_{e_H}(s_2)} < \frac{f_{e_L}(s_1)}{f_{e_L}(s_2)} \quad \overset{MLRP}{\Rightarrow} \quad e_L > e_H$$

This is a contradiction to the assumption. Hence there is no equilibrium with $e_H > e_L$.

ii) Now assume $e_H = e_L$. Then for $s > s^*$ the average productivity $w(s) = \bar{\theta}$ has to be offered because of perfect competition among the firms. Hence low types choose $e_L = 1$. Then $e_H = 1$ by assumption. So $e_H = e_L = 1$ is an equilibrium for $\bar{\theta} \geq v_H$. The wage schedule then satisfies $w(s) = \bar{\theta}$ for $s > s^*$ and $w(s) < v_L$ otherwise. ■

Proof of Proposition 2.

According to the theorem, $e_L > e_H$ in any such equilibrium. Then by MLRP for all $s_1 > s_2 > s^*$:

$$\frac{f_{e_H}(s_1)}{f_{e_H}(s_2)} < \frac{f_{e_L}(s_1)}{f_{e_L}(s_2)} \quad \Leftrightarrow \quad \frac{f_{e_H}(s_1)}{f_{e_L}(s_1)} < \frac{f_{e_H}(s_2)}{f_{e_L}(s_2)} \quad \Rightarrow w(s_1) < w(s_2)\,.$$

Remember $w(s) = 0$ for all $s \leq s^*$. Define $h : S \longrightarrow [0, v_H]$ as the wage that would be justified, if both types accepted at signal s:

$$h(s) := \frac{n f_{e_H}(s)\,\theta_H + f_{e_L}(s)\,\theta_L}{n f_{e_H}(s) + f_{e_L}(s)}$$

Due to the MLRP $h(s)$ is decreasing in s for $s > s^*$. Therefore if $h(\overline{s}) > v_H$, the first case in Proposition 2 holds. Otherwise, the second case holds for

$$s^{**} = \begin{cases} s^* & h(s) \leq v_H \ \forall s \\ \min\{s \in S : h(s) \leq v_H\} & \text{otherwise} \end{cases}$$

which is well defined. ■

Proof of Proposition 3.

Define $X = \{(e_L, e_H)|\, e_L \geq e_H\}$ and consider the correspondence $k : X \rightarrow [0,1] \times [0,1]$ with $k(e_L, e_H) = (e_L^*(e_L, e_H), e_H^*(e_L, e_H))$, where e_L^* and e_H^* are the optimal effort choices for individual low and high types given that everybody else chooses e_L or e_H according to their type. For the purpose of this proof, in order to deal with both cases from Proposition 2 at the same time, we write the first case like the second with $s^{**} = \overline{s} + 1 \notin S$ with $f_e(\overline{s}+1) = 0$ for all $e \in [0,1]$. Under this convention, recall that e_L and e_H determine $s^{**}(e_L, e_H)$ as in the proof of Proposition 2, and that the wage schedule $w_{e_L,e_H}(s)$ dictated by e_L and e_H is decreasing on $\{s \in S : s^{**}(e_L, e_H) > s > s^*\}$.

First show that $k : X \rightarrow [0,1] \times [0,1]$ is upper hemicontinuous (uhc). To see this, recall that $f_e(s)$ is continuous in e for all s. Hence, $w_{e_L,e_H}(s)$ is continuous in e_L and e_H for all $s^* < s < s^{**}(e_L, e_H)$. Therefore

$$U_i(w_{e_L,e_H}, e) = F_e(s^*)\, v_i + \int_{\{s \in S : s > s^*\}} \max(w_{e_L,e_H}(s), v_i)\, f_e(s)\, ds$$
$$= F_e(s^*)\, v_i + \int_{\{s \in S : s^{**}(e_L,e_H) > s > s^*\}} w_{e_L,e_H}(s)\, f_e(s)\, ds + (1 - F_e(s^{**}))\max(\theta_L, v_i)$$

is continuous in e, e_L and e_H for $i \in \{L, H\}$. Consequently, $e_i^*(e_L, e_H) = \underset{e \in [0,1]}{\arg\max}\, U_i(w, e)|_{e_L,e_H}$ is uhc according to the Theorem of the Maximum.

Next, we need to show that $e_L^* \geq e_H^*$, which implies $k : X \rightarrow X$. Then, k has a fixed point in X by Kakutani's fixed-point theorem, which establishes the claim.

Assume to the contrary, that one of the following holds:

i) $e_L > e_H$ and $e_H^* > e_L^*$. Then, $s^{**} > s^*$ has to hold, because otherwise $e_H^* = 0$. Define

$g_e(s) := \dfrac{f_e(s)}{F_e(s^{**}) - F_e(s^*)}$. Because high types behave optimally and in case of indifference choose the lower effort by assumption,

$$
\begin{aligned}
U_H(e_H^*) > U_H(e_L^*) &\Leftrightarrow \int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) f_{e_H^*}(s)\, ds > \int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) f_{e_L^*}(s)\, ds \\
&\Leftrightarrow \left(F_{e_H^*}(s^{**}) - F_{e_H^*}(s^*)\right) \int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) g_{e_H^*}(s)\, ds \qquad (*) \\
&> \left(F_{e_L^*}(s^{**}) - F_{e_L^*}(s^*)\right) \int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) g_{e_L^*}(s)\, ds
\end{aligned}
$$

Note that, by construction, $\int_{\{s\in S: s^{**}>s>s^*\}} g_e(s)\, ds = 1$ for all e, and that MLRP for f_e implies MLRP for g_e. Because $w(s)$ is decreasing on $\{s \in S : s^{**} > s > s^*\}$ it follows immediately that $\int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) g_{e_H^*}(s)\, ds < \int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) g_{e_L^*}(s)\, ds$. We then conclude from $(*)$, that $\left(F_{e_H^*}(s^{**}) - F_{e_H^*}(s^*)\right) > \left(F_{e_L^*}(s^{**}) - F_{e_L^*}(s^*)\right)$. Further note that $1 - F_{e_H^*}(s^{**}) > 1 - F_{e_L^*}(s^{**})$ due to MLRP. Remember $v_H > v_L$ and $\theta_L > v_L$.

With all this in mind consider low types' utility from playing e_H^* and e_L^* respectively:

$$
\begin{gathered}
U_L(e_H^*) - v_L = \\
\int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) f_{e_H^*}(s)\, ds + \left(F_{e_H^*}(s^{**}) - F_{e_H^*}(s^*)\right)(v_H - v_L) + \left(1 - F_{e_H^*}(s^{**})\right)(\theta_L - v_L) \\
> \int_{\{s\in S: s^{**}>s>s^*\}} (w(s) - v_H) f_{e_L^*}(s)\, ds + \left(F_{e_L^*}(s^{**}) - F_{e_L^*}(s^*)\right)(v_H - v_L) + \left(1 - F_{e_L^*}(s^{**})\right)(\theta_L - v_L) \\
= U_L(e_L^*) - v_L
\end{gathered}
$$

As e_H^* gives strictly higher utility to low types than e_L^*, e_L^* cannot be low types' best response to (e_L, e_H). This contradicts the assumption.

ii) $e_L = e_H$ and $e_H^* > e_L^*$. This implies $w(s)$ is constant for all $s > s^*$. If $w(s) \leq v_H$, high types choose $e_H = 0$. But then $w(s) = \theta_L > v_L$ and low types choose $e_L = 1$, hence $e_L^* > e_H^*$. If, instead, $w(s) > v_H$, then high types accept all wage offers for $s > s^*$. So $w(s) = \bar{\theta} > v_H$ must hold and both types choose $e_L = e_H = 1$. This is an equilibrium if and only if $\bar{\theta} > v_H$.

From Cases i) and ii) we conclude that indeed $k : X \to X$ and by Kakutani's fixed-point theorem, an equilibrium with $e_L^* \geq e_H^*$ always exists. The argument under Case (ii) establishes that an equilibrium with $e_L^* = e_H^*$ exists if and only if $\bar{\theta} > v_H$. ■

Proof of the claim in Example 3.

For expositional clarity we normalize $\theta_L \equiv 1$ and $v_L \equiv 0$.[14] If skilled high types accept

[14] Formally we have only one degree of freedom left for normalization, as we already set the productivity

only w_1 and not w_2, then $w_1 > w_2 = \theta_L$ has to hold. Clearly $\theta_L < v_H$ and $\theta_H > v_H$ are necessary and $e_H = 1/2$ will result. To verify whether a combination of θ_H and v_H allows for such an equilibrium, express θ_H and v_H as functions of e_L: The equilibrium exists if and only if the wage offers satisfy

$$w_1 = \frac{2\left(\frac{1}{4}\theta_H + e_L\left(1 - e_L\right)\right)}{2\left(\frac{1}{4} + e_L\left(1 - e_L\right)\right)} > v_H \tag{1}$$

to make high types accept w_1 and

$$w_2 = \frac{\frac{1}{4}\theta_H + e_L^2}{\frac{1}{4} + e_L^2} \leq v_H \tag{2}$$

to prevent high types from accepting w_2, because otherwise a firm could offer this wage and all high types would accept. Equivalently:

$$\frac{1}{4}\left(\theta_H - v_H\right) \overset{!}{>} e_L\left(1 - e_L\right)\left(v_H - 1\right) \tag{1'}$$

$$\frac{1}{4}\left(\theta_H - v_H\right) \overset{!}{\leq} e_L^2\left(v_H - 1\right) \tag{2'}$$

For $v_H > \bar{\theta} = \dfrac{\theta_H + 1}{2}$, $(2')$ holds trivially and for $e_L \geq \widehat{e}_L$ with

$$\widehat{e}_L = \frac{1}{2}\left(1 + \sqrt{1 - \frac{\theta_H - v_H}{v_H - 1}}\right)$$

$(1')$ holds, too.

The expected utility of low types is $U_L = 2e_L\left(1 - e_L\right)w_1 + e_L^2$. Consequently low types choose e_L according to the first order condition

$$\frac{\partial U_L}{\partial e_L} = \left(2 - 4e_L\right)w_1 + 2e_L = \left(2 - 4e_L\right)\frac{\frac{\theta_H}{4} + e_L\left(1 - e_L\right)}{\frac{1}{4} + e_L\left(1 - e_L\right)} + 2e_L \overset{!}{=} 0$$

of unskilled types to zero. Note, however, that setting it to any negative value would not change our results, so that the suggested normalization does not limit the generality of the proof.

From this condition we uniquely determine θ_H as a function of e_L:

$$\theta_H\left(e_L\right) = \frac{4e_L^3 - 8e_L^2 + 5e_L}{2e_L - 1}$$

Now the definition of $\widehat{e}_L$ yields an upper limit $\bar{v}\left(e_L\right)$ for values of v_H, below which the proposed equilibrium exists. Algebra yields

$$\bar{v}\left(e_L\right) = \frac{1 + \theta_H\left(e_L\right) - \left(2e_L - 1\right)^2}{2 - \left(2e_L - 1\right)^2} > \bar{\theta}$$

Considering $\theta_H\left(e_L\right)$ is continuous and monotonic in e_L on $\left]\frac{1}{2}, 1\right]$ with values $\theta_H \in [1, \infty[$, the inverse $e_L\left(\theta_H\right)$ exists. Hence for any $\theta_H \in [1, \infty[$ and $v_H \in \left]\bar{\theta}, \bar{v}\left(\theta_H\right)\right] \neq \emptyset$ one interior equilibrium has high types accept only w_1 and choose $e_H = 1/2$ and low types choose $e_L\left(\theta_H\right)$. The resulting wages are $w_1\left(e_L, \theta_H\right)$ and $w_2 = \theta_L = 1$ as zero profit dictates. ■

Proof of Proposition 4.
Define $h\left(s\right)$ as the wage offer justified if both types accept at signal $s > s^*$:

$$h\left(s\right) := \frac{n f_{e_H}\left(s\right)\theta_H + f_{e_L}\left(s\right)\theta_L}{n f_{e_H}\left(s\right) + f_{e_L}\left(s\right)}.$$

Claim. Perfect pooling can occur if and only if $a = a^* := \dfrac{\bar{\theta} - v_H}{\bar{\theta} - v_L} < 1$ and $\bar{\theta} > v_H$.

Proof. Consider two cases:
i) $\bar{\theta} > v_H$. For perfect pooling $e_H = e_L = e^*$ for some e^* and consequently $w\left(s\right) = h\left(s\right) = \bar{\theta} > v_H$ for $s > s^*$. So high types always accept employment for $s > s^*$. This is an equilibrium if and only if

$$e_L = e^* = \arg\max_e \left(F_e\left(s^*\right)v_L + \left(1 - F_e\left(s^*\right)\right)\bar{\theta} - c\left(e\right)\right)$$
$$e_H = e^* = \arg\max_e \left(F_e\left(s^*\right)v_H + \left(1 - F_e\left(s^*\right)\right)\bar{\theta} - ac\left(e\right)\right)$$

First order conditions are

$$\left.\frac{\partial F_e(s^*)}{\partial e}\right|_{e=e^*}\left(v_L-\bar{\theta}\right)=c'(e^*)$$
$$\left.\frac{\partial F_e(s^*)}{\partial e}\right|_{e=e^*}\left(v_H-\bar{\theta}\right)=ac'(e^*)$$

determining $a^*=\dfrac{\bar{\theta}-v_H}{\bar{\theta}-v_L}<1$ as we claimed.
ii) $\bar{\theta}\le v_H$. Then, given $e_H=e_L=e^*$, high types will not benefit from employment and hence choose $e_H=0$. But $e_L>0$ always holds. So, no perfect pooling is possible. $\|$

Claim. There is $\bar{a}<1$ sufficiently large, such that for $a>\bar{a}$ overeagerness dominates.

Proof. Assume to the contrary, that for all $\bar{a}<1$ there is $a>\bar{a}$ for which there is an equilibrium with $e_H\ge e_L$. In such equilibrium the wage schedule is monotonic. It can be shown analogous to the proof of Proposition 2 that there is $s^{**}<\bar{s}$ such that

$$w(s)=\begin{cases}0 & s\le s^*\\ \theta_L & s^*<s<s^{**}\\ h(s) & s\ge s^{**}\end{cases}$$

Define $Y:=\{(e_L,e_H)|\,e_L\le e_H\}$, the uhc correspondence $\tilde{k}:Y\to[0,1]\times[0,1]$ analogous to $k:X\to[0,1]\times[0,1]$ in the proof of Proposition 3 and $(e_L^*,e_H^*):=\tilde{k}(e_L,e_H)$. We are searching for $\bar{a}$ such that $e_L^*>e_H^*$ for all $a>\bar{a}$ and for all $(e_L,e_H)\in Y$. This would rule out equilibria where $e_H\ge e_L$ and simultaneously establish that $\tilde{k}:Y\to Y$, so that a fixed point of $\tilde{k}$, which constitutes an equilibrium, exists. In equilibrium individual low and high types maximize respectively:

$$U_H(e)=v_H+(1-F_e(s^{**}))\int_{\{s\in S:s>s^{**}\}}(w(s)-v_H)\frac{f_e(s)}{1-F_e(s^{**})}ds-ac(e)$$
$$U_L(e)=U_H(e)+F_e(s^*)(v_L-v_H)+(F_e(s^{**})-F_e(s^*))(\theta_L-v_H)-(1-a)c(e)$$

We now establish bounds on e_H^* and s^{**}:

- Low types will at least get θ_L for $s>s^*$ and $w(s)$ is increasing, strictly increasing somewhere. Let

$$\underline{e}_L:=\arg\max_e\{(1-F_e(s^*))(\theta_L-v_L)-c(e)\}>0$$

We now argue that $e_L^* \geq \underline{e}_L$, which implies by assumption that $e_H^* > e_L^* \geq \underline{e}_L$.

To see why $e_L^* \geq \underline{e}_L$, write

$$U_L(e) = v_L + (1 - F_e(s^*))(\theta_L - v_L) + \int_{\{s \in S : s > s^{**}\}} (w(s) - \theta_L)\, df_e(s) - c(e)$$

Let $\xi(s) := \begin{cases} 0 & s \leq s^{**} \\ w(s) - \theta_L & s > s^{**} \end{cases}$. Since for $\underline{e}_L < 1$ it must be that

$$\left. \frac{\partial \{(1 - F_e(s^*))(\theta_L - v_L) - c(e)\}}{\partial e} \right|_{\underline{e}_L} = 0$$

we have

$$\left. \frac{\partial U_L(e)}{\partial e} \right|_{\underline{e}_L} = \left. \frac{\partial \int_S \xi(s)\, df_e(s)}{\partial e} \right|_{\underline{e}_L}$$

Finally, since f_e on S satisfies the MLRP, and since ξ is weakly increasing on S, $\frac{\partial \int_S \xi(s) df_e(s)}{\partial e} \geq 0$ for all $e \in (0,1)$. Hence, $e_L^* = \arg\max_e U_L(e) \geq \underline{e}_L$ as claimed.

- High types receive at most θ_H for $s > s^*$. Following an analogous argument to the previous one, we see that $e_H^* < \overline{e}_H$, where $\theta_H (1 - F_{\overline{e}_H}(s^*)) = ac(\overline{e}_H)$ if a solution exists for $\overline{e}_H \in (0,1)$ and $\overline{e}_H = 1$ otherwise. Note that, in the latter case, $c(\overline{e}_H) < \infty$.

- Since wage is bounded above by θ_H, high types will certainly not exert effort unless $(1 - F_e(s^{**}))(\theta_H - v_H) > ac(e_H^*) > ac(e_L^*)$. This implies $s^{**} < \tilde{s}$ where $(1 - F_{\overline{e}_H}(\tilde{s}))(\theta_H - v_H) = ac(\underline{e}_L)$.

Hence, $e_H^* \in [\underline{e}_L, \overline{e}_H]$ and $s^* < s^{**} < \tilde{s}$. Further $f_e(s)$ has full support for all e, so MLRP implies $\frac{\partial F_e(s)}{\partial e} < 0$ for all $s \in S \cap (0,1)$. Now consider

$$\frac{\partial (U_L(e) - U_H(e))}{\partial e} = -\frac{\partial F_e(s^{**})}{\partial e}(v_H - \theta_L) - \frac{\partial F_e(s^*)}{\partial e}(\theta_L - v_L) - (1 - a)\, c'(e)$$

With the definitions $\phi(s) := \min_{e \in [\underline{e}_L, \overline{e}_H]} \left(-\frac{\partial F_e(s)}{\partial e} \right) > 0$ and $\underline{\phi} := \min_{s \in \{s \in S : s^* < s < \tilde{s}\}} \phi(s) > 0$ we find

$$\begin{aligned} \frac{\partial (U_L(e) - U_H(e))}{\partial e} &\geq \underline{\phi}(v_H - \theta_L) + \phi(s^*)(\theta_L - v_L) - (1 - a) \max_{e \in [\underline{e}_L, \overline{e}_H]} c'(e) \\ &\geq \underline{\phi}(v_H - v_L) - (1 - a)\, c'(\overline{e}_H) \end{aligned}$$

for all $e \in [\underline{e}_L, \overline{e}_H]$. Therefore, if $a > \bar{a} := 1 - \dfrac{\phi(v_H - v_L)}{c'(\overline{e}_H)}$, then $\dfrac{\partial (U_L(e) - U_H(e))}{\partial e} > 0$ for all $e \in [\underline{e}_L, \overline{e}_H]$. Clearly $\bar{a} < 1$.

As $U_H(e)$ is maximized in e_H^*, it must be that $U_H(e_H^*) \geq U_H(e)$ for all $e \leq e_H^*$ and $\left.\dfrac{\partial (U_L(e) - U_H(e))}{\partial e}\right|_{e=e_H^*} > 0$, which implies $e_L^* > e_H^*$. This contradicts the initial assumption. Hence, for all $a > \bar{a}$ there is no equilibrium with $e_H^* \geq e_L^*$; that is, $\tilde{k} : Y \to Y$, so that a fixed point exists. Hence, overeagerness dominates for $a > \bar{a} < 1$. $\|$

Claim. If costs are significant enough in the sense that $c'(0) > \underline{c}'$ for an appropriate $\underline{c}'$, then there is $\underline{a} < \bar{a}$ sufficiently small such that for $a < \underline{a}$ the classical signaling effect dominates.

Proof. Define

$$\underline{c}' := \max_{e \in [0,1], s \in \{s \in S : s > s^*\}} \left\{ -\frac{\partial F_e}{\partial e}(s^*)(v_H - v_L) + \frac{\partial F_e}{\partial e}(s)(v_H - \theta_L) \right\}$$

and assume contrary to the claim, that for all a there is an equilibrium with $e_H \leq e_L$. In such equilibrium the wage schedule is as in Proposition 2: There is $s^{**} \in S \cup \{\overline{s} + 1\}$ such that

$$w(s) = \begin{cases} 0 & s \leq s^* \\ h(s) & s^* < s < s^{**} \\ \theta_L & s \geq s^{**} \end{cases}$$

where we allow $s^{**} = \overline{s} + 1$ for a more convenient way of writing Proposition 2, as in the proof of Proposition 3. In equilibrium individual low and high types maximize respectively:

$$\begin{aligned} U_L(e) = v_L + (F_e(s^{**}) - F_e(s^*)) \int_{\{s \in S : s^* < s < s^{**}\}} (w(s) - v_L) g_e(s) \, ds \\ + (1 - F_e(s^{**}))(\theta_L - v_L) - c(e) \\ U_H(e) = U_L(e) + \{(1 - F_e(s^{**}))(v_H - \theta_L) + F_e(s^*)(v_H - v_L) + (1 - a) c(e)\} \end{aligned}$$

By definition of $\underline{c}'$,

$$\frac{\partial (U_H(e) - U_L(e))}{\partial e} =$$

$$-\frac{\partial F_e(s^{**})}{\partial e}(v_H - \theta_L) + \frac{\partial F_e(s^*)}{\partial e}(v_H - v_L) + (1 - a) c'(e) > -\underline{c}' + (1 - a) c'(e)$$

Recall that $\frac{\partial^2 c(e)}{\partial e^2} > 0$, such that $c'(e) \geq c'(0)$ for all $e \in [0, 1]$. Let $\underline{a} := 1 - \dfrac{\underline{c}'}{c'(0)}$, where

$\overline{a} > 0$ if and only if $c'(0) > \underline{c}'$. Then, for $a < \underline{a}$, $\dfrac{\partial (U_H(e) - U_L(e))}{\partial e} > 0$ for all $e \in [0,1]$. Define X and the uhc correspondence $k : X \to [0,1] \times [0,1]$ as in the proof of Proposition 3 with $k(e_L, e_H) = (e_L^*(e_L, e_H), e_H^*(e_L, e_H))$. We know that $U_L(e)$ is maximized in e_L^*. Therefore $U_H(e_L) \geq U_H(e)$ for all $e \leq e_L^*$ and $\left.\dfrac{\partial (U_H(e) - U_L(e))}{\partial e}\right|_{e=e_L^*} > 0$, which implies $e_H^* > e_L^*$. This contradicts the initial assumption. We just established that for all $a < \underline{a}$ and $c'(0) > \underline{c}'$ we have $k : X \to \{(e_L, e_H) | \, e_L < e_H\}$. That is, an equilibrium exists and classic signaling dominates. $\|$ $\blacksquare$

References

Araujo, Aloisio; Daniel Gottlieb and Humberto Moreira (2007) "A Model of Mixed Signals with Applications to Countersignaling." *RAND Journal of Economics*, 38(4), 1020-1043.

Benabou, Roland and Jean Tirole (2006) "Incentives and Prosocial Behavior." *American Economic Review*, 96(5), 1652-1678.

Bertrand, Marianne; Dean Karlan; Sendhil Mullainathan; Eldar Shafir and Jon Zinman (2010) "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." *The Quarterly Journal of Economics*, 125(1), 263-305.

Blume, Andreas; Oliver Board and Kohei Kawamura (2007) "Noisy Talk." *Theoretical Economics,* 2, 395-440.

Chen, Ying (2011) "Perturbed Communication Games with Honest Senders and Naive Receivers." *Journal of Economic Theory,* 146, 401-424.

Chung, Kim-Sau and Peter Esö (2013) "Persuasion and Learning by Countersignaling." *Economics Letters,* 123(3), 487-491.

Clements, Matthew T. (2011) "Low Quality as a Signal of High Quality." *Economics,* Vol. 5, 2011-5.

Daley, Brendan and Brett Green (2014) "Market Signaling with Grades." *Journal of Economic Theory,* 151, 114-145.

Feltovich, Nick; Rich Harbaugh and Ted To (2002) "Too Cool for School? Signaling and Countersignaling." *RAND Journal of Economics*, 33(4), 630-649.

Hvide, Hans K. (2003) "Education and the Allocation of Talent." *Journal of Labor Economics*, 21(4), 945-976.

Orzach, Ram; Per Baltzer Overgaard and Yair Tauman (2005) "Modest advertising signals strength." *RAND Journal of Economics*, 33(2), 340-358.

Orzach, Ram and Yair Tauman (2005) "Strategic Dropouts." *Games and Economic Behavior*, 50(1), 79-88.

Spence, Michael (1974) "Market Signaling." Cambridge, MA. Harvard University Press.

— (1973) "Job Market Signaling." *Quarterly Journal of Economics,* 87, 355-374.

Teoh, Siew Hong and Chuan Yang Hwang (1991) "Nondisclosure and Adverse Disclosure as Signals of Firm Value." *The Review of Financial Studies*, 4(2), 283-313.

Weiss, Andrew (1995) "Human Capital vs. Signaling Explanation of Wages." *The Journal of Economic Perspectives* 9(4), 133-154.