Overeagerness*

Philipp Sadowski†

July 2013

Abstract

We capture the impression that high types may send lower signals than low types in order not to appear too desperate. In contrast to the counter-signaling literature, we require only a noisy one-dimensional signal, where very low signal manifestations force 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 a very low signal manifestation.

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.

If such a reaction is justified, we should find that sometimes the less competitive send stronger signals than the most competitive. Indeed, Clements (2011) points out that high quality products often come in cheap packaging. For example, premium beer does not have the more convenient screw off caps found on low-quality beer; and, in the United States, newspapers with high-quality content are sold in broadsheet format instead of the more

---

*I thank Roland Benabou, Avinash Dixit, Tymon Tatur, and Huseyin Yildrim for useful comments.

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

1See Weiss (1995) for a convincing attempt in the context of wages.
convenient tabloid format. 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. Furthermore Felovich, Harbaugh and To (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. Similarly, some luxury brands, like good wineries, hardly advertise at all.

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, and is therefore 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 have a higher opportunity cost of being employed than low types. For example, high types may not only be more productive when employed, but also when self employed. This assumption should have intuitive appeal.

Second, opportunity costs become relevant as very low signal manifestations leave individuals facing 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.

- There could be three types in the population, unskilled, low skilled and high skilled, where the large majority is unskilled, does not have any productive value for the company, and faces a prohibitive cost for sending a high signal. This is convincing if, for example, some basic skill has to have been acquired over time and independent of ability.

---

2In 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.

3A Cointreau slogan in Germany, rather paradoxically, reads “We are so good, we don’t need advertisement” (“Wir sind so gut, wir brauchen keine Werbung”).

Considering that unskilled individuals have no incentive to behave strategically, the above scenarios are equivalent. We formulate our model according to the last interpretation, but Section 2.1 gives an example that is in line with the first.

Third, we assume that individuals have only a noisy signal available to them. There is, then, some risk for low and high types to end up with a very low signal manifestation. This, too, seems to be a realistic assumption.

These 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.

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. This paper merely argues that the intuitive additional assumption of 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 very low signal manifestation as the higher signal. When effort is costly, this should also be the signal bearing the higher direct cost.

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 to high types, then the classical signaling effect compensates over eagerness. 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. In particular, our model applies to situ-

\[\text{Consider the example of advertising: Clearly it is easier to highlight the positive features of a high quality product. For example consumer tests can be cited, etc. So, the same level of advertising should cost less money for high quality producers. At the same time, a high quality producer could make better use of every dollar he does not invest in a public advertising campaign, say by investing in the relations with his existing customer base, the same base that constitutes his better outside option to a successful campaign. Therefore, taking opportunity costs into account, it could even be that advertising is cheaper for a low type.}\]
ations where cheap talk is possible for both types. Accounting for type dependent costs of 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, Overgaard, and Tauman (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, Gottlieb, and Maureira (2004) address the observation that wages are non-monotonic in the GED of high school dropouts. Feltovich, Harbaugh, and To (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 (2013) 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 (2007) 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. Our work differs from those papers in sticking to a setup with one period and a one-dimensional noisy signal.\(^5\)

Section two 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 three considers a binomial distribution of signals as an instructive example. Section four concludes.

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 \(\theta_H > \theta_L\), and unskilled types

\(^5\)Due to the noise there can be partially separating equilibria, allowing us to meet the incentive constraint with a one-dimensional signal.
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 \( n \).

Applicants have a noisy and public signal with support in \([0, 1]\) available to them. We assume for now that it is cost free. Applicants without the required skill send signals from a distribution with support \([0, s^*]\), where \( s^* < 1 \). Let \( f_u(s) \) denote the according density function. Applicants with the required skill can choose the effort \( e \in [0, 1] \) they put into sending the signal, which will then be a random draw according to the density function \( f_e(s) \), where \( f_e(s) \) has full support for all \( e \). Let \( f_e(s) \) satisfy the Monotone Likelihood Ratio Property (MLRP): \( \forall s_1 > s_2 \) the ratio \( \frac{f_e(s_1)}{f_e(s_2)} \) is increasing in \( e \). Also let \( f_e(s) \) be continuous in \( e \) for all \( s \).

For \( i \in \{L, H\} \) the utility of a representative applicant is \( u_i(w | \text{accept offer}) = w \) from accepting employment at a certain wage \( w \) or \( u_i(w | \text{do not accept offer}) = v_i \) 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. If indifferent between accepting and declining an offer, applicants are assumed to accept. The wage schedule offered by firms is \( w(s), w : [0, 1] \rightarrow [0, \theta_H] \).

A strategy is an effort choice \( e \). A pure strategy Nash equilibrium then consists of effort choices \((e_L, e_H)\), where the resulting wages for \( s \in (s^*, 1] \) are

\[
    w(s) = \frac{nf_{eH}(s) \theta_L + f_{eL}(s) \theta_L}{nf_{eH}(s) + f_{eL}(s)}
\]

if high types accept such an offer and \( w(s) = \theta_L \) if they would decline it,\(^6\) and where applicants maximize their expected utility

\[
    U_i(w, e) = \int_0^1 \max(w(s), v_i) f_e(s) \, ds
\]

For \( s \in [0, s^*] \) wages are specified accordingly, taking into account the presence of un-

\(^6\)Formally firms could be considered as players, too, but due to the full support of the signal distribution, there is no out of equilibrium signal value. Due to zero profit they can, then, be understood as a mechanism that rewards applicants.
skilled 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 $w(s) < v_L < v_H$ for $s \in [0, s^\ast]$ has to hold in equilibrium.\footnote{Specifically, it has to be true that high types want to deviate from accepting a wage justified by both types accepting, and low types want to deviate from accepting a wage justified by only low types accepting it. Suppose, for simplicity, that unskilled types face opportunity cost $v_u = 0$ and draw a signal from the distribution $f_u$. Then}

$U_i(w, e) = F_e(s^\ast) v_i + \int_{s^\ast}^1 \max(w(s), v_i) f_e(s) ds$

where $F_e$ denotes the cumulative distribution function corresponding to $f_e$.

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

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.

**Proof.** 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 $s_2$:

$$\frac{nf_{eH}(s_1) \theta_H + f_{eL}(s_1) \theta_L}{nf_{eH}(s_1) + f_{eL}(s_1)} < \frac{nf_{eH}(s_2) \theta_H + f_{eL}(s_2) \theta_L}{nf_{eH}(s_2) + f_{eL}(s_2)}$$

for all $e_H, e_L \in [0, 1]$ and $s \in [0, s^\ast]$ is a sufficient constraint on $r$ to guarantee this.
which is equivalent to
\[
\frac{f_{e_H}(s_1)}{f_{e_L}(s_1)} < \frac{f_{e_H}(s_2)}{f_{e_L}(s_2)} \iff \frac{f_{e_H}(s_1)}{f_{e_L}(s_1)} < \frac{f_{e_H}(s_2)}{f_{e_L}(s_2)} \stackrel{\text{MLRP}}{\Rightarrow} e_L > e_H
\]

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 \( s \in (s^*, 1] \) the average productivity \( w(s) = \bar{\theta} := \frac{n\theta_H + \theta_L}{n + 1} \) 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. ■

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

**Proposition 2.** In every equilibrium with \( e_H < 1 \) there is \( s^{**} \geq s^* \) such that
\[
w(s) = \begin{cases} 
< v_L & s \leq s^* \\
\frac{n f_{e_H}(s) \theta_H + f_{e_L}(s) \theta_L}{n f_{e_H}(s) + f_{e_L}(s)} & s^* < s \leq s^{**} \\
\theta_L & s > s^{**}
\end{cases}
\]

**Proof.** According to the theorem, \( e_L > e_H \) in any such equilibrium. Then by MLRP for all \( s_1 > s_2 > s^* \):
\[
\frac{f_{e_H}(s_1)}{f_{e_L}(s_1)} < \frac{f_{e_H}(s_2)}{f_{e_L}(s_2)} \iff \frac{f_{e_H}(s_1)}{f_{e_L}(s_1)} < \frac{f_{e_H}(s_2)}{f_{e_L}(s_2)} \Rightarrow w(s_1) < w(s_2).
\]

Remember \( w(s) = 0 \) for all \( s \leq s^* \). Define \( h(s) \) as the wage that would be justified, if both types accepted at signal \( s \):
\[
h(s) := \frac{n f_{e_H}(s) \theta_H + f_{e_L}(s) \theta_L}{n f_{e_H}(s) + f_{e_L}(s)}
\]

Since \( f_e(s) \) is continuous in \( s \in (0, 1) \) for all \( e \), so is \( h(s) \). Due to the MLRP it is decreasing in \( s \). Therefore \( h^{-1}(.) \) is well defined. Then
\[
s^{**} = \begin{cases} 
1 & h(1) \geq v_H \\
s^* & h(s) < v_H \forall s \\
h^{-1}(v_H) & \text{otherwise}
\end{cases}
\]

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

The proposition implies that for $\bar{\theta} < v_H$ an equilibrium with $e_L > e_H$ exists. Section 3 demonstrates that it can also exist for $v_H > \bar{\theta}$. The proof is in the appendix.

The analysis of signaling models à 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. 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.

2.1. An Example: 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 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, $\theta_L$. 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, $\theta_L$ 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

\footnote{In addition there may be direct cost, as considered in section 2.2.}
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 a rare field experiment conducted by Bertrand, Karlan, Mullainathan, Shafir and Zinman (2010) in South Africa. 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.

2.2. Type dependent cost of effort

To this point cheap talk is possible in our setup. Clearly, a more general model allows for a type-dependent cost of effort for sending a certain signal, for example, preparation for an exam.

Signaling models typically assume the cost of sending a specific signal to be negatively
correlated with productivity $\theta$. Indeed, it seems reasonable to assume that high types would need less time for preparing a certain exam or, in general, send a particular signal. In our model, however, they also have a better outside option. Hence, the opportunity cost for spending a given amount of time is 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, 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.\footnote{See footnote 5 to the Introduction for a more explicit example in the context of advertising.}

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 $c(e), c : [0, 1] \rightarrow [0, \infty]$ 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.

**Theorem 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 $c'(0) > \bar{c}'$ for an appropriate $\bar{c}' > 0$, then there is $\bar{a} > 0$ sufficiently small such that for $a < \bar{a}$ the classical signaling effect dominates.

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

The proof is in the appendix. Theorem 4 immediately implies $\underline{a} \leq \bar{a}$ with strict inequality for $\bar{\theta} \geq v_H$.

### 3. Binomially Distributed Signals

We now consider a sender-receiver game as in section 2, where we specify the uncertainty: Applicants have access to a noisy, public and cost free signal. It will be called a test result. For simplicity, suppose the test consists 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 $s \in \{0, 1, 2\}$ counts the correct answers to the two questions.
considered. For unskilled applicants, none of the questions can ever be answered correctly. Both types of skilled applicants can choose the rate at which they answer question correctly. Let \( e_L, e_H \in [0, 1] \) be the rates chosen by the respective types. Test signals are then distributed binomially for skilled applicants. The probability not to get any question right (\( s = 0 \)) is \( p_0 = (1 - e)^2 \), the probability to get one of the questions right (\( s = 1 \)) is \( p_1 = 2(1 - e)e \) and the probability to get both questions right (\( s = 2 \)) is \( p_2 = e^2 \). For unskilled applicants the probability of \( s = 0 \) is \( p_0 = 1 \).

The wage schedule can condition on the value of \( s \), where \( w = (w_0, w_1, w_2) \) is the vector of conditional wage offers. Let there be sufficiently many unskilled applicants of low type such that \( w_0 < v_L \) has to hold as in section 2. Then, the expected utility faced by an applicant of type \( i \in \{L, H\} \) is

\[
U_i(w, e) = (1 - e)^2 v_i + 2e(1 - e) \max \{w_1, v_i\} + e^2 \max \{w_2, v_i\}.
\]

A strategy consists of an effort choice \( e \). A pure strategy Nash equilibrium then consists of effort choices \((e_L, e_H)\) and wages \( w = (w_0, w_1, w_2) \). Due to perfect competition on the firm’s side

\[
w_1(e_H, e_L) = \frac{ne_H(1 - e_H)\theta_H + e_L(1 - e_L)\theta_L}{ne_H(1 - e_H) + e_L(1 - e_L)}, \quad w_2(e_H, e_L) = \frac{ne_H^2\theta_H + e_L^2\theta_L}{ne_H^2 + e_L^2}
\]

are the wages justified by the average productivity of applicants with one or two correct answers \((s = 1 \text{ or } s = 2)\) respectively, if the offer is accepted by both types. The formulas make it clear that, as before, equilibria could be only partially separating.\(^{10}\)

In case type \( i \in \{L, H\} \) accept neither \( w_1 \) nor \( w_2 \), we assume \( e_i = 0 \) for tie-breaking.\(^{11}\) If \( i \) accepts only one wage offer, \( w_s \), and is indifferent to declining it, we assume that \( i \) accepts and maximizes the probability of corresponding outcome \( s \). So only equilibria in which all players of one type act the same are possible: In case of indifference between strategies their behavior is unambiguous. Mixing between \( e_i = 0 \) and \( e_i = 1 \) is also precluded. Zero profit wages given \( s \) are, then, either uniquely determined by the ratio of high types over low types with signal manifestation \( s \), or they are \( \theta_L \). Consequently optimal efforts \( e_H \) and \( e_L \) are also uniquely determined by utility maximization, as the expected utility is concave in the individual effort choice. Hence, no type \( i \in \{L, H\} \) can be mixing over effort choices \( e_i \in (0, 1) \). In sum, we can concentrate on finding equilibria in pure strategies.

\(^{10}\) Out-of-equilibrium beliefs on the firm’s side can now be relevant only for boundary cases with \( e_H, e_L \in \{0, 1\} \), where signals do not have full support. We will specify these beliefs in the proof of the fact below, which deals with these cases.

\(^{11}\) A very small cost of effort would make this the only equilibrium choice.
We gave up the assumption of full support of the distribution of the signal (in $e \in \{0, 1\}$ we have $s = 0$ or $s = 1$ with certainty, respectively), and thus we convince ourselves (in the appendix) that the following versions of Theorem 1 and Proposition 2 remain true:

**Theorem 5.** In any pure strategy Nash equilibrium of the model described above $1 \geq e_L > e_H \geq 0$ or $e_H = e_L = 1$.

The case $e_H = e_L = 1$ is equivalent to the equilibrium we would get for a binary signal $s \in \{0, 1\}$ where skilled individuals of low and high type are not distinguishable. Then, high types will work for $\tilde{\theta} := \frac{n\theta_H + \theta_L}{n + 1} \geq v_H$. This will be called “benchmark case” in the discussion below.

**Proposition 6.** In every equilibrium where $0 < e_H < 1$, $w_1 > w_2$ has to hold.

We now analyze the different equilibria. For expositional clarity and for the remainder of this section, normalize $\theta_L \equiv 1$, $v_L \equiv 0$. First, note that low types always work for $w_1, w_2 \geq 0$. For $e_H \neq 1$ we established that $w_1 > w_2$. So, high types either do not work at all or they accept to work only when offered $w_1$ or they are willing to work for both, $w_1$ and $w_2$. By assumption $\theta_H > \theta_L = 1$, thus consider $(v_H, \theta_H) \in [0, \infty) \times (1, \infty)$. For ease of notation also fix the ratio of high over low types to be $n = 1$; other values give analogous results. Call equilibria where high types never work “No participation” (N), those where they accept $w_1$ and $w_2$ “Full participation” (F), and the ones where they accept only $w_1$, call “Selective participation” (S).

Note that we can rule out fully separating equilibria, where the signals and the acceptance of a certain offer are a sufficient statistic for the type of the applicants and in which at least some high types accept employment. If they existed, high types would have to decline offer $w_2$. Hence, $e_H = 1/2$, $w_2 = \theta_L$. Also, low types would have to choose $e_L = 1$ in equilibrium, requiring $\frac{\partial U_L}{\partial e_L} \bigg|_{e_L=1} \geq 0$ where $U_L(e_L) = e_L^2 \theta_L + 2e_L (1 - e_L) \theta_H$ is the utility of an individual applicant of low type. So $\theta_L - \theta_H \geq 0$ would be necessary, which is precluded by assumption.

**Proposition 7.** For $n = 1$ and $\theta_H > 1$, and with parameter dependent boundaries $\bar{v}(\theta_H)$ and $\underline{v}(\theta_H)$ as given in the proof, a partially separating equilibrium exists for $v_H < \bar{v}(\theta_H)$. Specifically, the possible equilibria are:

- (N) High types never work, $e_L = 1$, $e_H = 0$, $w_1 \leq \theta_L$, and $w_2 = \theta_L = 1$ is the unique equilibrium for $v_H \in (\bar{v}(\theta_H), \infty)$. It exists for all $v_H > \theta_L = 1$.

\[12\text{Formally we have only one degree of freedom left for normalization, as we already set the productivity of unskilled types to zero. Note well, however, that setting it to any negative value would not change our results, so that the suggested normalization is valid.}\]
Figure 1: a) $e_L(\theta_H)$ for S-equilibria; b) Types of equilibria

- **(S)** High types accept only $w_1$, $e_H = 1/2$, $w_2 = \theta_L = 1$, and $e_L$ and $w_1$ are jointly determined by zero profit for the firm and the low type best responding. This exists as an equilibrium for $v_H \in [\underline{\nu}(\theta_H), \bar{\nu}(\theta_H)]$.

- **(F)** High types accept both, $w_1$ and $w_2$, $1 > e_L > e_H \geq \frac{1}{2}$, and $w_1 > w_2 > \theta_L = 1$ exists as an equilibrium for $v_H \in (0, \underline{\nu}(\theta_H)]$ and possibly also for $v_H \in (\underline{\nu}(\theta_H), \bar{\theta}]$.

- The benchmark equilibrium exists for $v_H \in (0, \bar{\theta}]$.

The proof is in the appendix. Figure a) visualizes this result. Figure b) shows the boundaries $\underline{\nu}(\theta_H)$, $\bar{\nu}(\theta_H)$ and $\bar{\theta}$ as a function of $\theta_H$. Recall that in the benchmark equilibrium all low and all high types with the required skill work and receive the same wage. It is equivalent to the only equilibrium with working high types for a binary signal. Thus, for $v_H > \bar{\theta}$, high types would never work with only a binary signal, whereas for $\bar{\theta} < v_H \leq \bar{\nu}(\theta_H)$ in the S-equilibrium half the high types will work for $w_1$. This is a clear Pareto improvement, as low types also gain on average.

4. 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 who cares exclusively about employees, high types taking the risk of executing their outside option obviously impose a negative externality on low types, who are paid less 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, in the example of binomially distributed signals, 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 if there is noise in their own signal.

5. Appendix

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. Recall that $e_L$ and $e_H$ determine $s^*(e_L, e_H)$ as in Proposition 2, and that the wage schedule $w_{e_L, e_H}(s)$ dictated by $e_L$ and $e_H$ is decreasing on $[s^*, s^{**}]$.

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, $s^*(e_L, e_H)$ and $w_{e_L, e_H}(s)$ are continuous in $e_L$ and $e_H$ for all $s^* < s < s^{**}(e_L, e_H)$. Therefore

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

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

is continuous in $e$, $e_L$ and $e_H$ for $i \in \{L, H\}$. Consequently, $e^*_i(e_L, e_H) = \arg \max_{e \in [0, 1]} U_i(w, e)_{e_L, e_H}$ is uhc according to the Theorem of the Maximum.
Next, we need to show that $e^*_L \geq e^*_H$, which implies $k : X \to X$. 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 $g_e(s) := \frac{f_e(s)}{F_e(s^{**}) - F_e(s^*)}$. Because high types behave optimally,

$$U_H(e^*_H \geq U_H(e^*_L) \iff \int_{s^*}^{s^{**}} (w(s) - v_H) f_e^*_H(s) ds \geq \int_{s^*}^{s^{**}} (w(s) - v_H) f_e^*_L(s) ds$$

$$\iff (F_e^*(s^{**}) - F_e^*(s^*)) \int_{s^*}^{s^{**}} (w(s) - v_H) g_e^*_H(s) ds$$

$$\geq (F_e^*_L(s^{**}) - F_e^*_L(s^*)) \int_{s^*}^{s^{**}} (w(s) - v_H) g_e^*_L(s) ds (*)$$

Note that, by construction, $\int_{s^*}^{s^{**}} g_e(s) ds = 1$ for all $e$, and that MLRP for $f_e$ implies MLRP for $g_e$. Because $w(s)$ is decreasing on $[s^*, s^{**}]$ it follows immediately that $\int_{s^*}^{s^{**}} (w(s) - v_H) g_e^*_H(s) ds < \int_{s^*}^{s^{**}} (w(s) - v_H) g_e^*_L(s) ds$. We then conclude from $(*)$, that $(F_e^*(s^{**}) - F_e^*(s^*)) > (F_e^*_L(s^{**}) - F_e^*_L(s^*))$. Further note that $1 - F_e^*(s^{**}) > 1 - F_e^*_L(s^{**})$ 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:

$$U_L(e^*_H) - v_L =$$

$$\int_{s^*}^{s^{**}} (w(s) - v_H) f_e^*_H(s) ds + (F_e^*(s^{**}) - F_e^*_L(s^*)) (v_H - v_L) + (1 - F_e^*(s^{**})) (\theta_L - v_L)$$

$$> \int_{s^*}^{s^{**}} (w(s) - v_H) f_e^*_L(s) ds + (F_e^*_L(s^{**}) - F_e^*_L(s^*)) (v_H - v_L) + (1 - F_e^*_L(s^{**})) (\theta_L - v_L)$$

$$= U_L(e^*_L) - v_L$$

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 $w(s) < v_H$, high types will never accept the offer and 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, $w(s) \geq v_H$, then high types accept all wage offers for $s > s^*$. So $w(s) = \bar{\theta} \geq v_H$ must hold and both types choose $e_L = e_H = 1$. This is an equilibrium if and only if $\bar{\theta} \geq v_H$.

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 2 establishes that an equilibrium with $e^*_L = e^*_H$ exists if and only if $\bar{\theta} \geq v_H$. ■
Proof of Theorem 4.
Define \( h(s) \) as the wage offer justified if both types accept at signal \( s > s^* \):

\[
h(s) := \frac{n f_{eH}(s) \theta_H + f_{eL}(s) \theta_L}{n f_{eH}(s) + f_{eL}(s)}.
\]

Claim. Perfect pooling can occur if and only if \( a = a^* := \frac{\tilde{\theta} - v_H}{\tilde{\theta} - v_L} < 1 \) and \( \tilde{\theta} \geq v_H \).

Proof. Consider two cases:

i) \( \tilde{\theta} \geq v_H \). For perfect pooling \( e_H = e_L = e^* \) for some \( e^* \) and consequently \( w(s) = h(s) = \tilde{\theta} \geq v_H \) for \( s > s^* \). So high types always accept employment for \( s > s^* \). This is an equilibrium if and only if

\[
e_L = e^* = \arg \max_e \left( F_e(s^*) v_L + (1 - F_e(s^*)) \tilde{\theta} - c(e) \right)
\]

\[
e_H = e^* = \arg \max_e \left( F_e(s^*) v_H + (1 - F_e(s^*)) \tilde{\theta} - ac(e) \right)
\]

First order conditions are

\[
\frac{\partial F_e(s^*)}{\partial e} \bigg|_{e=e^*} (v_L - \tilde{\theta}) = c'(e^*)
\]

\[
\frac{\partial F_e(s^*)}{\partial e} \bigg|_{e=e^*} (v_H - \tilde{\theta}) = ac'(e^*)
\]

determining \( a^* = \frac{\tilde{\theta} - v_H}{\tilde{\theta} - v_L} < 1 \) as we claimed.

ii) \( \tilde{\theta} < v_H \). Then, given \( e_H = e_L = e^* \), high types will never accept employment. Hence \( e_H = 0 \). 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 \( a \) there is an equilibrium with \( e_H \geq e_L \). In such equilibrium the wage schedule is monotonic. It can be shown analogous to the proof of
Proposition 2 that there is $s^{**}$ such that

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

Define $Y := \{(e_L, e_H) | e_L \leq e_H\}$, 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 $\tilde{a}$ such that $e^*_L > e^*_H$ for all $a > \tilde{a}$ and for all $(e_L, e_H) \in Y$. This would rule out equilibria where $e_H \geq e_L$ 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:

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

We now establish bounds on $e^*_H$ and $s^{**}$:

- Low types will at least get $\theta_L$ for $s > s^*$ and $w(s)$ is increasing. Therefore,

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

We set out under the assumption $e^*_H > e^*_L$.

- High types receive at most $\theta_H$ for $s > s^*$. Therefore $e^*_H < \bar{v}_H$, where $\theta_H (1 - F_{\bar{v}_H}(s^*)) = ac(\bar{v}_H)$ if a solution exists for $\bar{v}_H \in (0, 1)$ and $\bar{v}_H = 1$ otherwise. Not that, in the latter case, $c(\bar{v}_H) < \infty$.

- High types will certainly not exert effort unless $(1 - F_e(s^{**})) (\theta_H - v_H) \geq ac(e^*_H) > ac(e^*_s)$. This implies $s^{**} \leq \bar{s}$ where $(1 - F_{\bar{v}_H}(\bar{s})) (\theta_H - v_H) = ac(e_L)$.

Hence, $e^*_H \in [e_L, \bar{v}_H]$ and $s^{**} \in [s^*, \bar{s}]$. Further $f_e(s)$ has full support for all $e$, so MLRP implies $\frac{\partial F_e(s)}{\partial e} < 0$ for all $s \in (0, 1)$. 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 \( \phi (s) := \min_{e \in [e_L, e_H]} \left( -\frac{\partial F_e (s)}{\partial e} \right) > 0 \) and \( \hat{\phi} := \min_{s \in [s^*, 1]} \phi (s) > 0 \) we find

\[
\frac{\partial (U_L (e) - U_H (e))}{\partial e} \geq \hat{\phi} (v_H - \theta_L) + \phi (s^*) (\theta_L - v_L) - (1 - a) \max_{e \in [e_L, e_H]} c' (e)
\]

\[
\geq \hat{\phi} (v_H - v_L) - (1 - a) c' (\bar{c}_H)
\]

for all \( e \in [e_L, e_H] \). Therefore, if \( a > \bar{a} := 1 - \frac{\hat{\phi} (v_H - v_L)}{c' (\bar{c}_H)} \), then \( \frac{\partial (U_L (e) - U_H (e))}{\partial e} > 0 \) for all \( e \in [e_L, e_H] \). 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 \( \frac{\partial (U_L (e) - U_H (e))}{\partial e} \bigg|_{e = e_H^*} > 0 \), which implies \( e_L^* > e_H^* \). This contradicts the initial assumption. Hence, for all \( a \) there is no equilibrium with \( e_H^* \geq e_L^* \); that is, \( \bar{k} : Y \rightarrow Y \), so that a fixed point exists. Hence, overeagerness dominates for \( a > \bar{a} < 1 \). ||

**Claim.** If costs are significant enough in the sense that \( c_0 (0) > c_0 (1) \) for an appropriate \( c_0 \), then there is \( a < \bar{a} \) sufficiently small such that for \( a < \bar{a} \) the classical signaling effect dominates.

**Proof.** Define

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

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 \( s^{**} \) such that

\[
w(s) = \begin{cases} \\
0 & s \leq s^* \\
h(s) & s^* < s \leq s^{**} \\
\theta_L & s > s^{**} \\
\end{cases}
\]

In equilibrium individual low and high types maximize respectively:

\[
U_L (e) = v_L + (F_e (s^{**}) - F_e (s^*)) \int_{s^*}^{s^{**}} (w (s) - v_L) g_e (s) ds + (1 - F_e (s^{**})) (\theta_L - v_L) - c (e)
\]

\[
U_H (e) = U_L (e) + \{(1 - F_e (s^{**})) (v_H - \theta_L) + F_e (s^*) (v_H - v_L) + (1 - a) c (e)\}
\]

18
By definition of $c'$,
\[
\frac{\partial (U_H(e) - U_L(e))}{\partial e} = -\frac{\partial F_e(s^*)}{\partial e} (v_H - \theta_L) + \frac{\partial F_e(s^*)}{\partial e} (v_H - v_L) + (1 - a) c'(e) > -c' + (1 - a) c'(e)
\]
Recall that $\frac{\partial^2 c(e)}{\partial e^2} > 0$, such that $c'(e) > c'(0)$ for all $e \in [0, 1]$. Let $a := 1 - \frac{c'}{c'(0)}$, where $a > 0$ if and only if $c'(0) > c'$. Then, for $a < a$, $\frac{\partial (U_H(e) - U_L(e))}{\partial e} > 0$ 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 2 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 $\left. \frac{\partial (U_H(e) - U_L(e))}{\partial e} \right|_{e = e_L^*} > 0$, which implies $e_H^* > e_L^*$. This contradicts the initial assumption. We just established that for all $a < a$ and $c'(0) > c'$ we have $k : X \to \{(e_L, e_H) \mid e_L < e_H\}$. That is, an equilibrium exists and classic signaling dominates. \(\blacksquare\)

**Proof of Theorem 5.**

Consider two cases:

**i)** Suppose $e_H > e_L$. High types would have to accept at least $w_1$ or $w_2$, as otherwise $e_H = 0$. But then zero profit implies that $w_2 > w_1$ so that $w_2$ is surely accepted by both types. This implies low types maximize expected utility by choosing $e_L = 1$. As effort is bounded above by 1, we have $e_H \leq e_L$, which is a contradiction to the assumption.

**ii)** $e_H = e_L$. We know $e_L \neq 0$, thus high types have to accept at least one offer to justify $e_H \neq 0$. Then $w_1 = w_2$ would have to be offered for $e_L \neq 1$ because of perfect competition among the firms. Hence, low types would deviate and choose $e_L = 1$. Therefore, $e_H = e_L = 1$. This is an equilibrium if $w_2 = \frac{n\theta_H + \theta_L}{n + 1}$ and $w_1 \leq w_2$. \(\blacksquare\)

**Proof of Proposition 6.**

In such equilibrium we have $e_L > e_H$ according to Theorem 5, and high types accept at least one of the offers. Then zero profit dictates $w_1 > w_2$. \(\blacksquare\)

**Proof of Proposition 7.**

We first establish a fact about boundary equilibria that also implies the last part of the proposition.

**Fact.** There are two possible boundary equilibria: The benchmark equilibrium $e_H = e_L = 1$
exists for $v_H \leq \bar{\theta}$ and $e_H = 0, e_L = 1$ exists for $v_H > \theta_L$.

**Proof.** There are four potential kinds of equilibria where either $e_L \in \{0, 1\}$ or $e_H \in \{0, 1\}$. With $z \in [0, 1]$:

i) $e_L = z, e_H = 1$. This is the first equilibrium in the proposition for $z = 1$ and $v_H \leq \bar{\theta}$: The wage $w_1$ is only relevant out of equilibrium, because nobody ever ends up with $s = 1$. It can have any value $w_1 \leq w_2 = \bar{\theta}$ to support this equilibrium. So firms can believe that types are “equally likely” to send the out-of-equilibrium signal $s = 1$. If they thought it was more likely to come from a high type, both types would deviate from $e = 1$. For $v_H > \bar{\theta}$ high types will deviate. For $z \neq 1$ low types will deviate by choosing $e_L = 1$.

ii) $e_L = z, e_H = 0$ is the other boundary equilibrium in the proposition for $z = 1$ and $v_H > \theta_L = 1$: $w_2 = \theta_L = 1$ will not induce high types to deviate and $w_1$ is not determined, as nobody ever ends up with $s = 1$. It can have any value $w_1 \leq w_2 = 1$ to support this equilibrium. So here, firms have to believe $s = 1$ to come from a low type out of equilibrium. For $v_H < \theta_L = 1$ high types will deviate. For $z \neq 1$ low types deviate by choosing $e_L = 1$. For $v_H > \theta_L$ this equilibrium exists for any out-of-equilibrium beliefs.

iii) $e_L = 0, e_H = z$ has both types deviate by choosing $e = 1$.

iv) $e_L = 1, e_H = z$ can be the equilibrium described in i) for $z = 1$ and the one in ii) for $z = 0$. For $z \in ]0, 1/2[$ high types will deviate to $e_H = 0$ if $w_1 < v_H$ and to $e_H \geq 1/2$ for $w_1 \geq v_H$. Similarly, $z \geq 1/2$ implies $w_1 > w_2$, which makes low types deviate to benefit from $w_1$, considering their utility maximizing effort choice satisfies $e_L = \frac{w_1}{2w_1 - w_2}$.

We now take the different types of equilibria in turn.

S) In S-equilibria high types accept $w_1$ and not $w_2$. Thus, $w_1 > w_2 = \theta_L$ has to hold. Clearly $\theta_L < v_H$ and $\theta_H > v_H$ are necessary and $e_H = 1/2$ will result. To find all combinations of $\theta_H$ and $v_H$ that allow for S-equilibria, express $\theta_H$ and $v_H$ as functions of $e_L$: An S-equilibrium exists if and only if the wage offers satisfy

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

(1)

to make high types accept $w_1$ and

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

(2)
to prevent high types from accepting $w_2$, because otherwise a firm could offer this wage and all high types would accept. Equivalently:

$$\frac{1}{4} (\theta_H - v_H) \geq e_L (1 - e_L) (v_H - 1) \quad (1')$$

$$\frac{1}{4} (\theta_H - v_H) \leq e_L^2 (v_H - 1) \quad (2')$$

For $v_H > \bar{\theta} = \frac{\theta_H + 1}{2}$ the benchmark equilibrium exists. Then (2') holds trivially and for $e_L \geq e_L^{(1)}$ with

$$e_L^{(1)} = \frac{1}{2} \left( 1 + \sqrt{1 - \frac{\theta_H - v_H}{v_H - 1}} \right)$$

(1') holds, too. Similarly (1') holds trivially for $v_H \leq \bar{\theta}$ and for $e_L > e_L^{(2)}$ with

$$e_L^{(2)} = \frac{1}{2} \sqrt{\frac{\theta_H - v_H}{v_H - 1}}$$

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

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

From this condition we uniquely determine $\theta_H$ as a function of $e_L$:

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

Now the definitions of $e_L^{(1)}$ and $e_L^{(2)}$ yield an upper limit $\bar{\theta} (e_L)$ and a lower limit $\underline{\theta} (e_L)$ for values of $v_H$, between which an S-equilibrium exists. The expressions are

$$\bar{\theta} (e_L) = \frac{1 + \theta_H (e_L) - (2e_L - 1)^2}{2 - (2e_L - 1)^2}, \quad \underline{\theta} (e_L) = \frac{\theta_H (e_L) - 4e_L^2}{1 - 4e_L^2}$$

Considering $\theta_H (e_L)$ is continuous and monotonic in $e_L$ on $\left[\frac{1}{2}, 1\right]$ with values $\theta_H \in [1, \infty[$, the inverse $e_L (\theta_H)$ exists. Hence for any $\theta_H \in [1, \infty]$ and $v_H \in [\underline{\theta} (\theta_H), \bar{\theta} (\theta_H)]$ one interior equilibrium has high types accept only $w_1$ and choose $e_H = 1/2$ and low types choose $e_L (\theta_H)$. The resulting wages are $w_1 (e_L, \theta_H)$ and $w_2 = \theta_L = 1$ as zero profit dictates.
N) If $v_H > \bar{v} (e_L)$, then high types are not willing to work for $w_1 (e_L, \theta_H)$ as specified above. Because $e_H = e_L = 1$ is not an equilibrium for $\bar{v} < v_H$, the unique equilibrium implied by the assumptions is $e_L = 1, w_2 = \theta_L = 1, w_1 \leq \theta_L, e_H = 0$. Existence of the equilibrium for $v_H > \theta_L = 1$ is established by the fact.

F) If $v_H \leq v (\theta_H)$, then high types would deviate from an S-equilibrium by accepting the wage $w_2$ that would be justified if all high types accepted it. We now investigate whether an F-equilibrium exists for $v_H \leq v (\theta_H)$. If it does, it involves $e_L > e_H \geq 1/2$. Note that $v (\theta_H) < \bar{v}$ for $e_L > 1/2$. This guarantees $w_1 > v_H$. As before, define the best response of an individual in reaction to everybody else playing according to $e_L, e_H$ as $e_L^* (e_L, e_H)$ and $e_H^* (e_L, e_H)$ for low and high types respectively. Further, define the collections of fixed points of these best response correspondences as $\hat{e}_L (e_H) = \{ e_L \mid (e_L = e_L^* (e_L, e_H)) \}$ and $\hat{e}_H (e_L) = \{ e_H \mid (e_H = e_H^* (e_L, e_H)) \}$, respectively. Clearly, in equilibrium $e_L \in \hat{e}_L (e_H)$ and $e_H \in \hat{e}_H (e_L)$ have to hold. With these assumptions and definitions the following lemmata hold.

Lemma 8. In the above situation, for $v_H \leq v (\theta_H)$ the set $\hat{e}_L (e_H)$ is a function (it contains one element) that is increasing in $e_H$.

Proof. The first order condition for low types gives $e_L^* (e_L, e_H) = \frac{w_1 (e_L, e_H)}{2w_1 (e_L, e_H) - w_2 (e_L, e_H)}$. Note that, slightly abusing notation,

$$w_2 (e_L, e_H) = I_{w_2 \geq v_H} \frac{e_H^2 \theta_H + e_L^2}{e_H^2 + e_L^2} + I_{w_2 < v_H} \frac{e_H^2 \theta_H + e_L^2}{e_H^2 + e_L^2}$$

as the wage $w_2$ depends on whether high types accept it or not, which they do due to $v_H \leq v (\theta_H)$. Also, $w_1 > v_H$ is guaranteed by the assumptions. Considering $e_L > e_H \geq 1/2$, we have $w_1 > w_2$, $e_L^* \in \left( \frac{1}{2}, 1 \right]$. It is continuous in $e_L$, because $w_1$ and $w_2$ are. With $\frac{\partial w_1}{\partial e_L} > 0$, $\frac{\partial w_2}{\partial e_L} < 0$, we conclude

$$\frac{\partial e_L^*}{\partial e_L} = \frac{w_1 \frac{\partial w_2}{\partial e_L} - w_2 \frac{\partial w_1}{\partial e_L}}{(2w_1 - w_2)^2} < 0$$

Hence, $|\hat{e}_L (e_H)| = 1$ (because, obviously, identity $(e_L)$ is continuous, strictly increasing and
takes all values in \( \left[ \frac{1}{2}, 1 \right] \). From \( \frac{\partial w_1}{\partial e_H} < 0, \frac{\partial w_2}{\partial e_H} > 0 \) we conclude

\[
\frac{\partial e^*_L}{\partial e_H} = \frac{w_1 \frac{\partial w_2}{\partial e_H} - w_2 \frac{\partial w_1}{\partial e_H}}{(2w_1 - w_2)^2} > 0
\]

establishing that \( e^*_L(e_H) \) must indeed increase in \( e_H \). ||

**Lemma 9.** i) In the above situation \( k(e_L, e_H) := (e^*_L(e_L, e_H), e^*_H(e_L, e_H)) \) is continuous for \( v_H \leq v(\theta_H) \).

ii) For the set \( X_\varepsilon \in \mathbb{R}^2 \),

\[
X_\varepsilon := \left[ \frac{1}{2}, \frac{\theta_H}{\theta_H + \varepsilon} \right] \times \left[ \frac{1}{2}, 1 \right] \bigcap \{(e_L, e_H) | w_1(e_L, e_H) \geq w_2(e_L, e_H) + \varepsilon \}
\]

there is \( \xi \) small enough such that \( k : X_\varepsilon \to X_\varepsilon \) for all \( \varepsilon \leq \xi \) for \( v_H \leq \bar{v} \).

**Proof.** i) For \( v_H \leq v(\theta_H) \), we know \( w_2 \geq v_H \), and then \( k \) is obviously continuous, because \( e^*_L(e_L, e_H) \) and \( e^*_H(e_L, e_H) \) are continuous in both arguments.

ii) If \( w_1 \geq w_2 + \varepsilon \) then \( e^*_L \leq \frac{w_1}{w_1 + \varepsilon} \) and with \( w_1 \leq \theta_H \) this implies \( e^*_L \leq \frac{\theta_H}{\theta_H + \varepsilon} \). Remember that \( e_L \geq 1/2 \) and \( e_H \geq 1/2 \) are guaranteed by the restrictions prior to the lemma. Let \( \delta \) satisfy \( w_2 = w_1 - \delta \) with \( \delta > 0 \). Then best responses are \( e^*_L = \frac{w_1}{w_1 + \delta} \) for low types and

\[
e^*_H = \begin{cases} 
\frac{w_1 - v_H}{w_1 + \delta - v_H} & \text{for } w_2 \geq v_H \\
\frac{1}{2} & \text{else}
\end{cases}
\]

for high types. Given \( e^*_L \) and \( e^*_H \), zero profit and straightforward algebra have the following implications for the corresponding wages \( w^*_1 \) and \( w^*_2 \):

- For \( w_2 \geq v_H \):

\[
w^*_1 - w^*_2 = \frac{(e_L - e_H)(\theta_H - 1)}{e^*_H (1 - e_L) + e^*_L (1 - e_H) + (e_H + e_L - 2e_H e_L)}
\]

Collecting terms,

\[
e_L - e_H = \frac{v_H \delta}{(w_1 + \delta)(w_1 + \delta - v_H)}, \quad 1 - e_H = \frac{\delta}{w_1 + \delta - v_H},
\]

23
$1 - e_L = \frac{\delta}{w_1 + \delta}$, and $e_H + e_L - 2e_H e_L = \frac{(2w_1 - v_H)\delta}{(w_1 + \delta)(w_1 + \delta - v_H)}$ gives

$$w^*_1 - w^*_2 = \frac{v_H(\theta_H - 1)}{(2w_1 - v_H) + \frac{(w_1 - v_H)^2(w_1 + \delta)^2}{w_1} + \frac{w_2^2(w_1 + \delta - v_H)^2}{(w_1 + \delta)^2(w_1 - v_H)}} =: \xi_1$$

Note that $\xi_1 > 0$.

- For $w_2 < v_H$ (and hence high types rejecting the offer $w_2$): $w^*_1 - w^*_2 = \frac{(\theta_H - 1)}{1 + 4e_L(1 - e_L)} =: \xi_2 > 0$.

Define $\xi := \min \{\xi_1, \xi_2\}$. Then $\xi > 0$. Choose $\varepsilon < \xi$ to establish that indeed $k : X_\varepsilon \to X_\varepsilon$ with $X_\varepsilon$ defined as in the Lemma. ||

**Corollary 1.** In the situation in which both lemmata apply, there is at least one equilibrium $(e_H, e_L)$ such that $e_H = e^*_H(e_L, e_H)$ and $e_L = e^*_L(e_L, e_H)$. If $(\tilde{e}_H, \tilde{e}_L)$ is another such equilibrium, then $\frac{e_L - \tilde{e}_L}{e_H - \tilde{e}_H} \geq 0$.

**Proof.** Lemma 2 guarantees the existence of a fixed point of $k$ in $X_\varepsilon$ for $\varepsilon$ small enough and $v_H \leq v(\theta_H)$ due to Brouwer’s fixed-point theorem. Any such fixed point is an equilibrium as specified in the corollary. According to Lemma 1, $\hat{e}_L(.)$ is an increasing function in $e_H$ and any equilibrium has to be on the graph of this function. This establishes the second claim of the corollary. ||

Now notice that $X_\varepsilon$ does not contain either boundary equilibrium. Also for $v_H \leq v(\theta_H)$, as noted above, any other equilibrium must involve $w_2 \geq v_H$. Therefore $e_H > 1/2$ must hold for high types optimal effort choice, where they accept both offers, $w_1$ and $w_2$. Hence an F-equilibrium exists in addition to the benchmark case. ■

**References**


Bertrand, Marianne; Dean Karlan; Sendhil Mullainathan; Eldar Shafir and Jon Zinman (2010) “What’s Advertising Content Worth? Evidence from a Consumer Credit Marketing...


