Signaling with Career Concerns*

Kim-Sau Chung
University of Minnesota, Department of Economics

Péter Eső
Northwestern University, Kellogg School, MEDS

August 2008

Abstract

We analyze a game where a player who is privately, but imperfectly, informed about the state of nature (his productivity) can choose among overt actions that generate public signals about the state. His payoff depends in the short run on the public perception of his productivity, and in the long run on the state and his career choice. Since learning the public signals is helpful for his long-run choice, the player faces a conflict between verifiable and costly signaling: Choosing a more-informative action can signal confidence in his productivity, but it also signals self-doubt as a high type can more easily forego learning about the state. We show that under certain conditions a tri-partite equilibrium exists where low and high types pool on a less-informative action while intermediate types send a more-informative public signal. This equilibrium is non-monotonic and the average productivity of an agent choosing a less informative action is greater. We show that both features are present in any separating equilibrium. We discuss applications such as signaling and countersignaling in talent markets.

JEL classification: D82, D86. Keywords: signaling, career concerns

*We thank seminar participants at the Hong Kong University of Science and Technology, Northwestern University, University if Hong Kong, University of Texas at Austin, Washington University St. Louis, SED 2005 in Budapest, CETC 2007 in Montréal, and in particular Yeon-Koo Che, Hanming Fang, Sven Feldman, Drew Fudenberg, Stephen Morris, Alessandro Pavan, Nicolas Sahuguet, Kane Sweeney, Jeroen Swinkels for comments, and Renato Gomes for comments and excellent research assistance.

1 Introduction

Many economic activities involve agents generating public information about their qualities; these signals are often *informative for the agents themselves* as well as the outside world. Drug companies carry out or pay for experiments in order to convince regulators and customers that their products are safe and effective. However, the outcome of a preclinical trial is also useful for the company to determine if more investment in the drug is worthwhile. Another often-studied example is that of individuals who join organizations, choose certain activities, or participate in higher education in order to reveal their abilities to employers (or the rest of society). Note, however, that the grades received at school are useful for the individual, too, in evaluating his or her career options.

In our game a player, who has imperfect private information about a payoff-relevant state of nature (his productivity), chooses among overt actions that generate additional public signals regarding the true state. His actions are ordered according to the associated public signal's informativeness regarding the state, but do not differ in terms of their direct costs. In the first period (the "short run"), after the agent's action and the generated public signal are observed, the market pays him his expected productivity. Then, in period 2 (the "long run"), the agent chooses between a constant payoff and a wage proportional to his true productivity, and the game ends. The first-period public signal regarding the state of nature informs not only the market but also the agent's second-period decision because it updates his beliefs about the state. However, the precision of this signal is determined by the agent's action, which, by being observed by the market, may also affect his first-period payoff. We look for sequential equilibria where the player chooses different actions depending on his information.

An interesting feature of this game is that choosing a more informative public signal has conflicting effects on the agent's payoff. If the agent privately knows that his productivity is likely to be high, then an informative action is attractive for him as it is more likely to generate a favorable public signal. Indeed, if there existed a costless action revealing the state with *certainty* (a free, "hard" signal), then the agent would be forced to play it in order to be perceived as a high type (Grossman (1981), Milgrom (1981)). However, if no action is perfectly informative, then a less-informative action can be a useful signaling device à la Spence (1974), because foregoing the long-run benefits of learning about his own productivity is less costly for an agent with a higher expected productivity. These two modes of signaling (via verifiable signals vs. costly effort) have been studied separately before, but it is their *interaction* that provides interesting results in our model.

We show that under certain conditions, this game has a *tri-partite separating equilibrium*, where the agent sends a more-informative public signal only when he is moderately optimistic about his productivity; when he has either a sufficiently high or a sufficiently low type, he chooses a less-informative action.[1] Although both high and low types choose a less-informative action, they do so for different reasons. Low types shy away from a more-informative action because they are "afraid of the truth". High types, in contrast, want to signal that they are so confident in their productivity that they have nothing to learn about themselves. Overall, we show that the choice of a less-informative action signals strength, not the agent's fear from the truth. In equilibrium, enough high types choose a more-informative action allowing the low types to hide behind them, which curtails the Milgrom and Grossman-type unraveling result.

There are two noteworthy properties of our tri-partite equilibria. First, the agent's equilibrium strategy is non-monotonic, hence the equilibrium is different from the kind of "threshold-equilibrium" that we often see in signaling games. Second, in equilibrium, the average productivity of the agent choosing a less informative action is higher than that of the agent choosing a more informative one. We show that both features are common to *any equilibrium* where not all types of the agent choose the same action.

The reason why a less informative action is associated with higher expected productivity is somewhat subtle. Clearly, a less-informative action is *costly* as it decreases the value of the player's option to choose the state-independent payoff in the long run. But low types of the agent also *benefit* from sending a less-informative public signal as it is less likely to reveal the state, which in their case is likely to be low. Nevertheless, an equilibrium where (on average) low types choose a less-informative action cannot exist. The key step is to show that if the market's beliefs regarding the average types choosing each action were equal, then this average type would get the same short-run payoff from either action, and higher types would gain more from the more-informative one. After figuring in the long-run option value, all types at or above the average would still strictly

[1] Unsurprisingly, there may also exist pooling equilibria. For example, a trivial equilibrium where the agent always chooses the more-informative action, regardless of his private information, always exists. In our analysis, we focus on informative (signaling) equilibria.

prefer the more-informative action. This contradicts the assumption that the market's beliefs (that the average types choosing either action are equal) are rational. In order to restore equilibrium, the market's expectation of the average type choosing the less-informative action must be raised.

The intution for the non-monotonicity of every signaling equilibrium is the following. As we argued, a less-informative action is perceived by the market as a signal of confidence, but it is also less likely to generate a favorable signal in case the agent's productivity is high, and it is less valuable for the agent for the purpose of learning about the state. This tradeoff disappears for the most pessimistic agent (the type that *knows* that his productivity is low)—he does not care about learning, and he is actually glad that the signal is less likely to reveal the state. Therefore the lowest type of the agent chooses the less-informative action. Hence the types choosing this action cannot be uniformly higher than the types playing the more-informative action.

For a concrete example that corresponds to our dynamic signaling game, consider a budding artist is uncertain, but not completely uninformed, about her talent. She can launch her career with either a "traditional" or an "experimental" project (the former could be a painting or sculpture, the latter an installation). The quality and reception of her first work is informative regarding her artistic talent. The artist's short-run payoff (the price of her piece or the prize she wins with it) is correlated with the market's expectation of her talent given the project choice and the public signal (buzz) generated by it. The success of her first work also informs the artist whether she should continue her career in art or become, say, a decorator at a department store. If she remains an artist then, in the long run, the world learns her talent and appreciates (pays) her accordingly; if she quits art then her payoff is independent of her talent. The crucial, but reasonable, assumption is that a traditional project gives a more informative signal regarding the artist's talent than an experimental one. This is so because it is difficult for art critics and the public to evaluate unusual works of art. The question is which project the artist should choose in equilibrium, and what the market can infer from that choice.

Our results imply that the average talent of a young artist choosing an experimental project for her debut is greater. However, young artists who choose non-traditional projects at the beginning of their careers are not uniformly more talented than those choosing traditional projects. In fact, according to our model, we would expect to find truly gifted and also utterly untalented individuals among those who choose "the road less traveled". This seems to agree with our casual observation of talent markets in art and other creative endeavors.

The model has other applications as well. For example, the agent can be the product manager at a pharmaceutical company planning to test the effectiveness of a new drug (which corresponds to the unknown state or "productivity"). The manager has private information about the drug, and he can choose either an in-house experiment, or to provide a grant to a university-affiliated research team. It is reasonable to assume that the signal generated by the outside investigators is more informative than that of an in-house test.[2] The manager's pay is tied to the company's stock price as long as he stays with the company. While in the long run, the company's stock price will reflect the drug's eventual effectiveness, in the short run it is set according to the market's expectation of the drug's quality based on both the type and outcome of the experiment. Again, the question is which experiment to choose. A more informative test would enable the manager to make a better decision whether or not to continue to pursue the drug, but the mere act of choosing it may have an adverse impact on the short-run stock price. Our results imply that a less-informative, in-house test would be an indication that the manager is either very optimistic or very pessimistic about the product's quality, but also that, on average, the drug's prospects are better than they would be had the manager opted for outside testing.

Mainstream explanations of signaling phenomena usually rely on variants of Spence's (1973) model.[3] The starting point is an adverse selection situation; in addition, the privately informed agent can engage in a certain costly activity interpreted as a "signal". The key assumption is that the signaling activity is relatively less costly for an agent that has higher quality. This sorting (or single-crossing) condition enables high-quality agents to separate themselves from low-quality ones by choosing a sufficiently high level of the signal so that imitation is not worthwhile.

It has been pointed out (the observation going back to Veblen (1899)) that in many signaling situations we only see intermediate types sending the costly

[2]Outsourcing the tests may also cost less than doing it in-house. Assuming that an action associated with a more-precise signal has a lower direct cost would not alter our results.

[3]For a textbook exposition, see Fudenberg and Tirole (1991), Chapter 7. For the earliest examples of signaling models, see Spence (1973) on education, Nelson (1974) on advertising, Ross (1977) on the choice of a firm's financial structure, and Zahavi (1975) on mate selection in the animal kingdom.

signals, while very high productivity agents seem not to engage in such activity. For example, college dropouts include some of the most talented (not to mention richest) members of society. Feltovich, Harbaugh and To (2002) cite other examples as well: the truly rich do not flaunt their wealth, only the "nouveau rich" do; a person of the highest character does not bother to disprove accusations, only people with average reputations do; and so on.

Feltovich, Harbaugh and To (2002) model these "countersignaling" phenomena in a variant of a Spencian signaling game where the market receives an additional stochastic signal about the agent's type besides observing his action. Under certain conditions, medium types find it worthwhile to differentiate themselves from low types by traditional wasteful signaling, while high types—confident that in the end the exogenous signal will separate them from the low types—can afford not to signal in the traditional sense. In the end, the market is able to distinguish all three types of the sender, which is not the case in the model we study.[4]

Our assumptions and results differ from those in this line of research in other ways as well. Ours is not a Spencian signaling game because it is not inherently cheaper for a higher type to choose a more informative action. This is so because the agent's short-run payoff depends on the market's beliefs about which types choose each action, while his long-run payoff depends on the value of learning about his productivity from the signal generated by his action. The latter "learning benefit" from a more informative action is small for very low and very high types (the ones that are almost sure about their productivity), therefore the cost of a less informative action is not even monotonic.[5] Moreover, in our model, *all* informative equilibria have the property that some low and some high types pool on an action different from the one chosen by intermediate types. In the modified Spencian models cited above, some type of equilibrium refinement is needed to get a similar prediction.

[4] Hvide (2003) proposes a model with two sectors for employment: one where the wage depends on talent, and one where it does not. An individual who is privately informed about his ability may enter either sector right away, or get more education (private signals about his talent) before making his choice. Education is relatively cheaper for more talented people. A fully separating equilibrium (whose existence depends on parameter values) is where low types enter the flat-wage sector, high types choose the talent-based sector, and medium types get more education before making a choice. This, too, is a Spencian signalling game, and in equilibrium all types separate.

[5] The single-crossing property cannot be re-established even by transforming the type space (i.e., by relabeling types). This will become clear as we describe the model.

There exist a few other papers that explain interesting phenomena by way of non-monotonic signaling. The structure of these models and the results derived from them are quite different from ours, however. Baliga and Sjöström (2004) study an arms race preceded by a two-message cheap talk stage. They derive an equilibrium in which weak and strong types pool on the "dove" message while intermediate types play "hawk" in the cheap-talk stage. These messages allow coordination between the weak and intermediate types in the arms race, while strong types, not interested in coordination, pool with the weak in order to surprise their opponent. A similar effect, called sandbagging, is obtained in the context of jump bidding in auctions by Hörner and Sahuguet (2007). They argue that in a two-stage, two-player, private-values auction with costly bidding, a bidder with a high valuation may initially bid low ("sandbag", showing weakness) in order to soften competition in the second round. A counteracting incentive for the high type is to jump-bid in the first round hoping to deter the other bidder's entry in the second stage. In their model, medium-valuation bidders jump bid with probability one, while high-valuation bidders mix between jump-bidding and sandbagging.

The inefficiency result of our model (i.e., some types of the agent endogenously choose an inefficient method to learn about their own talents) is related to the inefficiency result of Brandenburger and Polak (1996). In both models, the agent cares about not only his productivity in the long run, but the market's current perception of his future productivity as well. In our model, this "short-term reputational concern" distorts the agent's incentive to learn about his own talent; in Brandenburger and Polak's it induces a manager to make the decision that the market wants to see (instead of the decision that maximizes the firm's long term profitability). One crucial difference, however, is that in Brandenburger and Polak's model, the "short-term reputational concern" eliminates every possibility of separating equilibrium. In our model, separating equilibria are possible, and necessarily take a non-monotonic form.

While the agent's short-term reputational concern distorts his incentives to do the "right" thing, his long-term career concerns determine the manner in which these incentives are being distorted. Our model is hence a contribution to the literature on career concerns, which studies various implications of an agent's long-term career concerns on his short-term behavior. In Holmström (1999), an agent's career concerns help motivate him to exert effort, which otherwise cannot be rewarded with an enforceable incentive contract. In Morris (2001), an informed advisor, who otherwise would have current incentive to truthfully reveal her information to her advisee, may refrain from doing so because she is concerned of her long-term reputation as an unbiased advisor.

In Ottaviani and Sorensen (2006a,b), the advisor is concerned of her reputation as being accurate (instead of being unbiased), and this concern in turn reduces the credibility of her short-term advice, so much so that truthful revelation becomes impossible. In Prendergast and Stole (1996), career concerns have opposite effects on young and old investors. Young investors tend to exaggerate their reactions to new information in order to signal that they are fast learners. On the contrary, old investors are more conservative in order to signal that they have always been fast learners and hence have already learned enough in the past. In Avery and Chevalier (1999), young investors who know little about their own ability herd in their investment behavior as in Scharfstein and Stein (1990). But as they get older and learn more about their abilities, they choose to "anti-herd" in order to signal that they are confident in themselves.

Finally, our model is also marginally related to cheap-talk games. Such games are an extreme form of non-Spencian signaling games, where cost differentials across different actions (messages) are type-independent as all messages are costless. Nevertheless, separating equilibria are still possible, because the receiver's (or the market's) reactions to different messages are different, and this creates endogenous type-dependent cost differentials across messages. In a clever twist of the standard setup of cheap talk games, Fang (2001) allows those cost differentials to be stochastic, while maintaining the assumption that they are type-independent. Endogenous type-dependent cost differentials can arise as in standard cheap talk models, and separating equilbria exist where different actions result in different market reactions. Fang (2001) interprets these different actions as different cultural activities, and uses this model to explain why productivity-unrelated cultural activities would nevertheless be rewarded differently by the market.

The rest of the paper is structured as follows: We set up the model in Section 2, and perform a preliminary analysis of the payoffs in Section 3. We prove the existence of tri-partite equilibria in Section 4, and show that some of its properties generalize to all signaling equilibria in Section 5. We conclude in Section 6, and present omitted proofs in the Appendix.

2 The Model

In this section we formally describe our model of signaling with career concerns. First, a partially-informed, risk-neutral agent chooses among observable actions that generate public signals about his true productivity. The market observes his action and the signal generated by it, and pays him a wage equal to his expected productivity. Upon observing all this, the agent chooses between an additional fixed payoff and a payoff that is proportional to his true productivity.

Denote the agent's productivity (the unobservable state of the world) by ω, and assume that it can take one of two values, H (high) or L (low), $H > L$. The prior distribution of ω is commonly known. Before the game starts, the agent observes a private signal regarding the state of nature. The signal generates a posterior distribution of ω; indeed, without any loss of generality, we can identify the agent's private information with his updated belief that the state of nature is H. That is, the agent's type, denoted by θ, is simply $\theta = \Pr(\omega = H)$. From an outside observer's perspective, the agent's type is drawn according to a commonly known distribution F with full support on $[0, 1]$. The ex ante expectation of θ is simply the commonly known prior probability that the state of nature is H.

There are two periods, and for notational simplicity, no discounting. The agent is assumed to be risk neutral. In the first period, the agent undertakes a publicly observable action. In order to simplify the exposition we assume that there are two alternatives available to him, a_1 and a_2. (All our results go through with an arbitrary number of actions.) Each action generates a random signal conditional on ω that is observable to the agent and the market alike. The realization of the public signal is denoted by $y \in \{H, L\}$. The restrictions that y is binary and that realizations of y correspond to realizations of ω are imposed purely for convenience and do not affect the results. The distribution of y conditional on a_i is characterized by $\pi_i \equiv \Pr(y = \omega | \, \omega, a_i)$ for $i = 1, 2$. Without loss of generality, let $\pi_i \geq 1/2$ for $i = 1, 2$. Our key assumption is that action a_1 generates a *more informative signal* about ω than a_2 does, that is, $\pi_1 > \pi_2$. The parameters π_1 and π_2 are commonly known.

After action a_i and signal value y are publicly observed, the agent is paid the expectation of his true productivity (the expectation of ω) given all publicly available information, including a_i, y, and the agent's equilibrium strategy, $\theta \mapsto a(\theta)$. This wage can be thought of as a "credence wage" for the agent's first-period performance (or services), which the market values according to the agent's yet unobservable productivity. In our earlier example, the budding artist's debut project was rewarded by the market (art speculators) according to their expectation of the artist's talent given the type and quality of her first art piece.

In the second period, the agent again chooses between two actions, labeled "in" and "out". If he stays in then he gets a payoff proportional to his true productivity, ω. If he chooses to get out then he gets a fixed payment, K. One may interpret the second period as the "long run", and the agent's choice between "in" and "out" as the reduced form of some more complex continuation game: If the agent continues with his activity then his productivity is eventually learned by the market, and he gets rewarded accordingly. However, he can also choose an outside option whose value is independent of his talent.[6]

Denote the agent's second-period updated belief that his productivity is high (given that he knows θ and observes y generated by a_i) by θ_i^y, that is,

$$\theta_i^y(\theta) = \Pr\left(\omega = H \mid \theta, y, a_i\right). \tag{1}$$

Note that $\theta_i^L(\theta) \leq \theta \leq \theta_i^H(\theta)$ with $E_y\left[\theta_i^y(\theta)\right] = \theta$, that is, the second-period belief is a mean-preserving spread of the first period belief, θ. (The spread is wider if π_i is larger.) In the second period, the agent chooses "in" whenever θ_i^y exceeds a certain threshold that depends on the value of the outside option, K.

To summarize, the order of moves in the game and the payoffs are as follows.

0. Nature chooses $\theta \in [0, 1]$ according to c.d.f. F, and picks either $\omega = H$ or $\omega = L$ with probabilities θ and $(1 - \theta)$, respectively. The risk-neutral agent privately learns his type, θ, while his productivity, ω, remains unknown.

1. The agent chooses a publicly observable action from $\{a_1, a_2\}$. Nature generates a publicly observable signal y where $y = \omega$ with probability π_i for action a_i ($i = 1, 2$), and $\pi_1 > \pi_2 \geq 1/2$. The agent is paid a wage that equals $E[\omega|a(\cdot), a_i, y]$, his expected productivity given the equilibrium strategy, the action taken, and the signal generated by the action.

2. The agent chooses between staying "in" and getting "out". The former yields a payoff proportional to ω while the latter yields a payoff of K.

[6] Allowing K to depend on the agent's productivity would not alter our results, as long as the outside option is less sensitive to ω than the agent's payoff when he stays "in".

We focus on perfect Bayesian equilibria (see Fudenberg and Tirole (1991), Definition 8.1), which can also be called rational-expectations equilibria due to the fact that the second player is the "market". In equilibrium, in the second period, the agent chooses "in" if and only if $\theta_i^y(\theta)$, his posterior belief about the state being high (given his type, action, and the generated public signal), exceeds a certain threshold that depends on K. In the first period, an equilibrium is characterized by the agent's choice of action conditional on his type, $a(\theta) \in \{a_1, a_2\}$ for all $\theta \in [0,1]$, and the market's belief that the agent's productivity is high given his action, $x_i \in [0,1]$ for $i = 1, 2$. In equilibrium, the market's beliefs must be consistent with the agent's strategy, which in turn has to be an optimal choice for the agent given his type and the market's beliefs.

As we already mentioned, there are uninformative equilibria where all types of the player choose to send the same public signal. For example, pooling on the more-informative action is always an equilibrium, supported by the market's out-of-equilibrium beliefs that only very low type(s) would choose the less-informative action.

In what follows we analyze informative equilibria, that is, rational-expectations equilibria where in period 1 both actions are taken by some types of the agent. A tuple $\langle a(\cdot), x_1, x_2\rangle$ is called an informative (or separating) equilibrium if $a^{-1}(a_1)$ and $a^{-1}(a_2)$ are both non-empty, and

$$a(\theta) = \begin{cases} a_1 & \text{if } W_1(\theta, x_1) + T_1(\theta_1, K) > W_2(\theta, x_2) + T_2(\theta, K), \\ a_2 & \text{if } W_1(\theta, x_1) + T_1(\theta_1, K) < W_2(\theta, x_2) + T_2(\theta, K), \end{cases} \tag{2}$$

$$x_i = \Pr(\omega = H \mid a(\cdot), a_i). \tag{3}$$

The first condition requires that the agent choose his most preferred action with type θ given the market's beliefs and the payoff functions; the second condition states that the market's beliefs are rational given the agent's strategy.

3 A Preliminary Analysis of the Payoffs

In this section we derive certain useful properties of the agent's payoff function. These properties will be used in Sections 4 and 5, where informative equilibria are analyzed.

The agent's payoff (in expectation, at the beginning of the first period) consists of two terms: his expected wage in the first period, and his future expected payoff from being able to choose between "in" and "out" in period two. The first-period expected wage is a function of his type, the action that he chooses, and the market's belief about his productivity that is associated with the action. The agent's expectation at the beginning of the game of the "option value" that he will enjoy in the second period also depends on his type and the action that he chooses in period one, but it does not depend on the market's perception of his productivity based on his initial choice. We formally define and derive these two parts of the agent's total payoff in turn.

3.1 Payoffs in the first period

Recall that the market's belief (estimated probability) regarding $\omega = H$ when the agent takes action a_i is denoted by x_i, and that the agent's equilibrium strategy is denoted by $a : [0,1] \to \{a_1, a_2\}$. In what follows we normalize the agent's productivity levels so that his expected productivity coincides with the estimated probability that $\omega = H$, that is, we set $H = 1$ and $L = 0$. This is without any loss of generality because the transformation is affine and the agent is risk neutral.

Let $W_i(\theta, x_i)$ denote the agent's expected first-period wage with type θ when he takes action a_i associated with market belief x_i. (The probability that the action generates a signal equal to the agent's true productivity, π_i, is a parameter that is suppressed by this notation.) The first-period payoff, W_i, is determined as follows. First, given the agent's strategy and his chosen action, the market's updated (posterior) belief that the agent's productivity is high when the signal generated by his action is y can be calculated by Bayes' rule as

$$\Pr(\omega = H \mid y, a(\cdot), a_i) = \frac{\Pr(y \mid \omega = H, a(\cdot), a_i) \Pr(\omega = H \mid a(\cdot), a_i)}{\Pr(y \mid a(\cdot), a_i)}.$$

The market wage paid to the agent given his action choice, a_i, and the signal realization, y, is $w_i^y = \Pr(\omega = H \mid y, a(\cdot), a_i)$. Using the above equality,

$$w_i^H = \frac{\pi_i x_i}{\pi_i x_i + (1-\pi_i)(1-x_i)}, \tag{4}$$

$$w_i^L = \frac{(1-\pi_i) x_i}{(1-\pi_i) x_i + \pi_i (1-x_i)}. \tag{5}$$

Here w_i^H (respectively, w_i^L) is the wage that the agent receives when he chooses action a_i associated with market belief x_i and the publicly observed signal happens to be H (respectively, L). If $\pi_i = 1/2$ then $w_i^H = w_i^L = x_i$ because the signal does not provide any new information about ω. However, if $\pi_i > 1/2$ then the agent's wage is higher when the signal realization is higher, $w_i^H > w_i^L$.

$W_i(\theta, x_i)$ is the agent's expectation of his first-period wage given θ, that is,

$$W_i(\theta, x_i) = \theta \left[\pi_i w_i^H + (1-\pi_i) w_i^L\right] + (1-\theta)\left[(1-\pi_i) w_i^H + \pi_i w_i^L\right]. \quad (6)$$

By substituting in w_i^H and w_i^L from (4) and (5) into this equation and rearranging terms we get

$$W_i(\theta, x_i) = \frac{\pi_i \theta + (1-\pi_i)(1-\theta)}{\pi_i x_i + (1-\pi_i)(1-x_i)} \pi_i x_i + \frac{(1-\pi_i)\theta + \pi_i(1-\theta)}{(1-\pi_i) x_i + \pi_i (1-x_i)} (1-\pi_i) x_i. \quad (7)$$

Notice that the agent's expected first-period wage is affine in θ, his initial belief regarding his productivity. We summarize other useful properties of $W_i(\theta, x_i)$ in the following lemma. Figure 1 illustrates W_i graphically.

Lemma 1 *If $\pi_i = 1/2$ then $W_i(\theta, x_i) \equiv x_i$. If $\pi_i > 1/2$ then the agent's expected wage in period 1 satisfies:*

(i) $W_i(\theta, 0) \equiv 0$ and $W_i(\theta, 1) \equiv 1$.

(ii) For all $x_i \in (0,1)$, $W_i(\theta, x_i)$ is strictly increasing in θ and x_i.

(iii) For all x_i, $W_i(x_i, x_i) = x_i$.

(iv) For all $x_i \in (0,1)$, $W_i(\theta, x_i)$ is strictly increasing in π_i if $\theta > x_i$. Conversely, $W_i(\theta, x_i)$ is strictly decreasing in π_i if $\theta < x_i$.

Proof. See the Appendix. ■

Notice that holding the precision of the signal-generating action (π_i) fixed, both the intercept and the slope of the short-run wage depend on the market's beliefs regarding the average productivity of the agent that takes that action. In particular, if the market's expectations are low (x_i is low), W_i starts out low and has a small slope. As we increase x_i, the expected talent associated with action a_i, the short-run expected wage increases and becomes more sensitive to the agent's private information. However, as the market's belief approaches certainty in the agent's high productivity, the wage becomes less and less sensitive to θ.

The comparison of first-period expected wage schedules resulting from different actions is difficult because the intercepts and slopes of the W_i functions depend on the relative precisions of the two signals (i.e., π_1 and π_2), and also the market's beliefs regarding the talent of the agent taking the different actions (i.e., x_1 and x_2). Since the market's beliefs are endogenous in the model, not

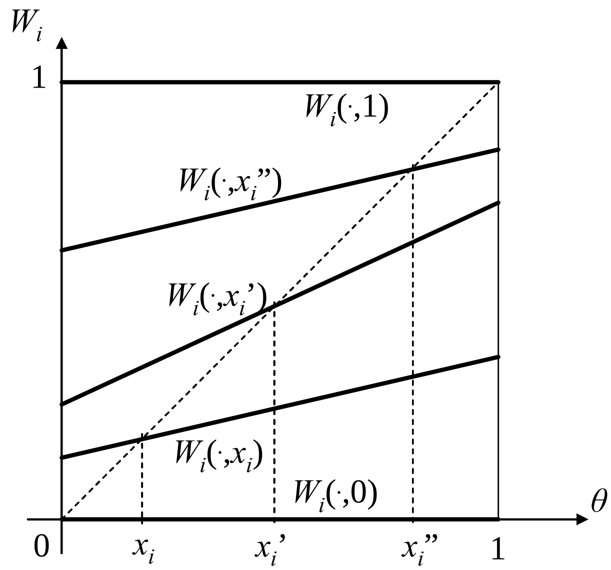

Figure 1: The first-period expected wage.

much can be said in advance regarding the difference between W_1 and W_2 at a particular θ. For example, $W_1(\cdot, x_1)$ can intersect $W_2(\cdot, x_2)$ from below or above (but only once), if the two graphs intersect at all.

By part (iii) of Lemma 1, if $x_1 = x_2 = x$ (i.e., the market's expectation of the agent's productivity is the same for both actions), then no matter how precise the actions are, the expected first period wage of type $\theta = x$ does not depend on the action choice, that is, $W_1(\theta, x) = W_2(\theta, x) = x$ for $\theta = x$. From part (iv) of Lemma 1 we also know that the expected first period wage (W_i) is increasing in the precision of the signal (π_i) if and only if the agent's type is greater than the market's expectation ($\theta > x_i$). A more precise signal is beneficial for the agent in the short run only if his type is better than the average type that chooses it.[7]

3.2 Payoffs in the second period

Now we turn to the characterization of the agent's second-period payoff.

Let $T_i(\theta, K)$ denote the agent's expectation at the beginning of period 1 of his

[7] From parts (ii) and (iv) of Lemma 1 it also follows that if $x_1 = x_2 = x$ but $\pi_1 > \pi_2$ then $W_1(\theta, x)$ crosses the 45 degree line at $\theta = x$ steeper than $W_2(\theta, x)$ does.

benefit from the second-period option to choose between getting a constant payoff K and a payoff equal to his true productivity. (Again, the parameter π_i is implicit in our notation.) Recall that $\theta_i^y(\theta)$, defined in equation (1), denotes the agent's updated (posterior) belief at the beginning of period 2 that his productivity is high given that action a_i generated signal y, and that his prior belief was θ. In particular, by Bayes' rule,

$$\theta_i^L(\theta) = \frac{(1-\pi_i)\theta}{(1-\pi_i)\theta + \pi_i(1-\theta)}, \tag{8}$$

$$\theta_i^H(\theta) = \frac{\pi_i\theta}{\pi_i\theta + (1-\pi_i)(1-\theta)}. \tag{9}$$

The property that observing y is informative for the agent regarding his productivity means that θ_i^y is a mean-preserving spread around θ, that is,

$$\Pr(y = L|\theta, a_i)\theta_i^L(\theta) + \Pr(y = H|\theta, a_i)\theta_i^H(\theta) \equiv \theta.$$

Since the agent chooses "in" over "out" in period 2 if and only if $\theta_i^y(\theta) \geq K$, the option value he gets from this choice is $\max\{\theta_i^y(\theta) - K, 0\}$. At the beginning of period 1, the agent does not know the realization of y yet, hence the expected value of his second-period option is

$$T_i(\theta, K) = E_y\left[\max\{\theta_i^y(\theta) - K, 0\} \mid \theta\right].$$

Figure 2 illustrates graphically the derivation and properties of $T_i(\theta, K)$. As it can be seen in the figure, $T_i(\theta, K) = 0$ for all θ such that $\theta_i^H(\theta) \leq K$, and $T_i(\theta, K) = \theta - K$ for all θ such that $K \leq \theta_i^L(\theta)$, and T_i is convex in θ.[8] Finally, for θ such that $\theta_i^L(\theta) < K < \theta_i^H(\theta)$, we have

$$\begin{aligned} T_i(\theta, K) &= \Pr(y = H \mid \theta, a_i)\left(\theta_i^H(\theta) - K\right) \\ &= \pi_i\theta - (\pi_i\theta + (1-\pi_i)(1-\theta))K. \end{aligned} \tag{10}$$

There is a difference between the second-period benefit generated by action a_1 and a_2 that arises as follows. Action a_1 is more informative than a_2, hence the agent's posterior beliefs are more spread out under a_1 than they are under a_2:

$$\theta_1^L(\theta) < \theta_2^L(\theta) \leq \theta \leq \theta_2^H(\theta) < \theta_1^H(\theta).$$

[8] It may be useful to note that similar qualitative properties would hold even if we had more than two possible realizations of y.

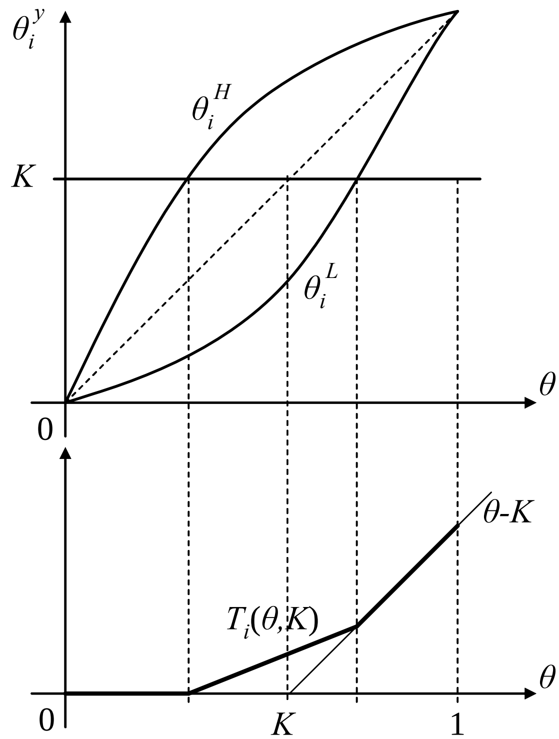

Figure 2: The second-period option value

Since the second-period option value, max $\{\theta_i^y - K, 0\}$, is convex in θ_i^y, the period 1 expectation of it is greater under action a_1 when θ_i^y is more spread out. That is, a more informative action generates a greater payoff in the second period because the agent learns more about his own productivity and so the value of the option to stay in or get out is greater.[9] The property that a_1 generates a weakly greater period 2 benefit than a_2 does *for all types of the agent* implies that a_2 is a costlier action compared to a_1.[10] However, the "cost" of action a_2 is not monotonic in the agent's type (in fact, it is zero at $\theta = 0$ and $\theta = 1$).

[9] This result holds as long as the player's indirect profit in the continuation game is a *convex function* of his updated second-period belief regarding his productivity.

[10] In Figure 2 we can easily see the "cost" of a totally uninformative action relative to action a_i whose second-period payoff T_i is depicted in the lower panel. Notice that the agent's payoff in period 2 after choosing an uninformative signal-generating action in period 1 is max $\{\theta - K, 0\}$. Therefore, the "cost" of choosing this action over a_i is $T_i(\theta, K) - \max\{\theta - K, 0\}$, which is zero near $\theta = 0$ and $\theta = 1$, and peaks at $\theta = K$.

The agent's total expected payoff from choosing action a_i in the first period (given that the market's belief associated with action a_i is x_i) is $W_i(\theta, x_i) + T_i(\theta, K)$. From the preceding analysis it is clear that our game is not, and cannot be transfomed into, a Spencian (monotonic) signaling game. In particular, the relative cost or benefit of choosing an action over the other is in part determined by the market's beliefs about the average type choosing each action, and the payoff-difference may be non-monotonic in the player's type.

4 Existence of Tri-Partite Equilibria

In this section we establish the existence of *tri-partite separating equilibria* where very low and very high types choose a less informative action, while intermediate type(s) choose a more informative action. Proving the existence of tri-partite equilibria involves a different machinery from proving the existence of equilibria in general. In particular, simple fixed-point arguments do not suffice, because there is no guarantee that the fixed points correspond to tri-partite equilibria instead of, say, pooling equilibria, which always exist.

We proceed with our analysis under two, alternative sets of conditions.

First, we consider a situation where the agent's type comes from a discrete distribution on exactly *three types*: low, medium, and high. We establish sufficient conditions under which an equilibrium exists where the low and high types pool on the less-informative action and the medium type chooses the more-informative one. The results for the three-type discrete distribution are useful for constructing examples and can also be used in applications.[11] The disadvantage of this approach is that the sufficient conditions for existence put joint restrictions on the three-type distribution and the relative informativeness of the public signals.

Second, and perhaps more interestingly, we consider type-distributions that are *continuous with a full support* on $[0, 1]$. Without making any additional assumptions on the distribution, we show that when the agent's actions are not too informative, an equilibrium exists that partitions the types of the agent in three increasing subsets: low types, medium types, and high types. Low and high types pool on the less-informative action, while medium types choose the more-informative one.

[11] Note that some of the countersignaling models in the literature, notably Feltovich, Harbaugh and To (2002), have a similar, three-type discrete structure as well.

4.1 Discrete type distributions

Suppose that the agent's type is distributed on three values, $\theta_L < \theta_M < \theta_H$, with probability weights (p_L, p_M, p_H).[12] We shall make joint assumptions on this three-type distribution and the informativeness of the public signals that guarantee the existence of a tri-partite equilibrium whenever the player's second-period state-independent payoff option (K) is in the neighborhood of θ_M. In that equilibrium, the intermediate type plays action a_1, while the extreme types pool on a_2. It is useful to introduce the notation

$$\mu_{HL} = \frac{p_L\theta_L + p_H\theta_H}{p_L + p_H},$$

the expected productivity of the player given that his type is either θ_L or θ_H.

Since the state-independent outside option, K, is in the neighborhood of θ_M, the medium type is the most likely one to gain from learning about the state. Indeed, we shall assume that the first-period public signal does not matter for the extreme types' choices between "in" and "out" in the second period. If $K \approx \theta_M$ then a set of simple sufficient conditions for this is

$$\theta_1^H(\theta_L) \equiv \frac{\pi_1\theta_L}{\pi_1\theta_L + (1-\pi_1)(1-\theta_L)} \leq \theta_M, \tag{11}$$

$$\theta_1^L(\theta_H) \equiv \frac{(1-\pi_1)\theta_H}{(1-\pi_1)\theta_H + \pi_1(1-\theta_H)} \geq \theta_M. \tag{12}$$

The two conditions mean that type θ_L chooses "out" in the second period even if the first-period signal is high, and type θ_H chooses "in" even if the realization of y is low. If conditions (11) and (12) hold for action a_1 then they also hold for a_2 because the latter action is less informative. Therefore, neither signal-generating action provides more option value for the extreme types; the less-informative action is "costless" for types θ_L and θ_H. As $\pi_1 \to 1/2$ the conditions (11)-(12) simplify to $\theta_L \leq \theta_M \leq \theta_H$; for $\pi_1 > 1/2$ the conditions essentially require that the three types be sufficiently spread out.

We shall prove the existence of a tri-partite equilibrium under two additional conditions. The first one is

$$W_1(1, \theta_M) \leq W_2(0, \mu_{HL}), \tag{13}$$

[12] In Section 2 we assumed the distribution of θ has full support on $[0,1]$. Therefore, the three-type "discrete" type distribution considered here is really an $(\varepsilon, 1-\varepsilon)$ mixture of (any) full-support type distribution and the distribution on $\{\theta_L, \theta_M, \theta_H\}$, where $\varepsilon > 0$ is arbitrarily small.

where W_i is defined by equation (7). This condition implies that the player gets a lower first-period wage choosing the more-informative action than the less-informative one, as long as the market believes that the medium type plays the former and the extreme types the latter. Since the right-hand side of (13) is increasing in μ_{HL}, the condition essentially puts a lower bound on the average productivity of the extreme types for a fixed θ_M, π_1 and π_2.

The second condition is

$$\mu_{HL} \leq \theta_M + 2(\pi_1 - \pi_2)\theta_M(1 - \theta_M). \tag{14}$$

In contrast to (13), this condition puts an upper bound on μ_{HL} given θ_M, π_1 and π_2. There exist parameter values that satisfy conditions (13) and (14); see the numerical example after Proposition 1 below.

The following proposition states that when K is sufficiently close to θ_M, the conditions (11)–(14) are sufficient for the existence of an equilibrium where θ_L and θ_H pool on action a_2 and θ_M plays a_1.

Proposition 1 *Suppose that the type distribution is discrete on $\theta_L < \theta_M < \theta_H$ with probability weights (p_L, p_M, p_H). If the inequalities (11)–(14) hold, then for K sufficiently close to θ_M, there exists an equilibrium where θ_L and θ_H choose a_2 and θ_M plays a_1 in the first period.*

Proof. In the proposed tri-partite equilibrium the market's beliefs are $x_1 = \theta_M$ and $x_2 = \mu_{HL}$. We now derive necessary-sufficient conditions for such an equilibrium and show that they hold under the hypothesis of the proposition.

By (11) and (12), if K is sufficiently close to θ_M, then the high and low types prefer a_2 over a_1 if and only if

$$W_1(\theta, \theta_M) \leq W_2(\theta, \mu_{HL}) \text{ for } \theta \in \{\theta_L, \theta_H\}. \tag{15}$$

This inequality is satisfied because W_1 and W_2 are both weakly increasing in θ and (13) holds by assumption. Note that by the linearity of W_i the inequality (15) also holds at $\theta = \theta_M$ (in fact, strictly). Since $W_i(\theta_M, \theta_M) = \theta_M$ for $i = 1, 2$ and $W_i(\theta, x_i)$ is strictly increasing in x_i (both by Lemma 1), $W_1(\theta_M, \theta_M) < W_2(\theta_M, \mu_{HL})$ implies $\mu_{HL} > \theta_M$.

The medium type prefers a_1 over a_2 if and only if

$$W_2(\theta_M, \mu_{HL}) + T_2(\theta_M, K) \leq W_1(\theta_M, \theta_M) + T_1(\theta_M, K). \tag{16}$$

By (10), for K sufficiently close to θ_M, the second-period payoff, $T_i(\theta_M, K)$, approximately equals $(2\pi_i - 1)\theta_M(1-\theta_M)$, and the previous inequality becomes $W_2(\theta_M, \mu_{HL}) \leq \theta_M + 2(\pi_1 - \pi_2)\theta_M(1-\theta_M)$. By $\mu_{HL} > \theta_M$ and $W_2(\theta_M, \theta_M) = \theta_M$, we have $W_2(\theta_M, \mu_{HL}) < \mu_{HL}$, therefore (14) is a sufficient condition for the strict version of (16).

The conditions of the proposed tri-partite equilibrium are (15)–(16), which are both implied by conditions (11)–(14). This completes the proof. ■

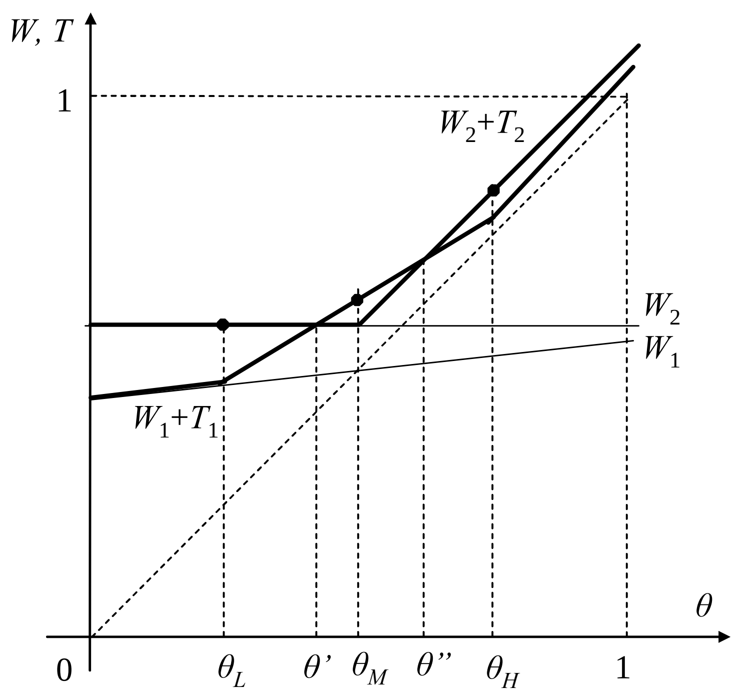

Figure 3: Tri-partite equilibrium in the discrete case

The construction of the equilibrium is illustrated in Figure 3, which corresponds to the following numerical example.

Example 1 *Let* $\pi_1 = 2/3$, $\pi_2 = 1/2$. *Asssume that the support of the type distribution is* $\theta_L = 1/3$, $\theta_M = K = 1/2$, $\theta_H = 2/3$, *and the probability weights* (p_L, p_M, p_H) *are such that* $x_2 \equiv (p_L\theta_L + p_H\theta_H)/(p_L + p_H) \in (5/9, 7/12)$, *for example,* $(p_L, p_M, p_H) = (7/48, 1/2, 17/48)$. *There exists a separating equilibrium where* θ_L *and* θ_H *pool on action* a_2 *while* θ_M *chooses* a_1.

In the example the first-period wage of the agent choosing action a_2 is $W_2 \equiv x_2 \in (5/9, 7/12)$ because the action is uninformative.[13] This wage exceeds the first-period wage from action a_1 for any type of the agent because $W_1(1, 1/2) = 5/9$ by equation (7) and the slope of W_1 is positive. This verifies condition (13). The parameters are chosen so that types θ_L and θ_H do not enjoy positive second-period option values from either first-period action, hence by $W_2 > W_1$ both θ_L and θ_H indeed strictly prefer action a_2 to a_1. Finally, type θ_M plays action a_1 because $W_1(\theta_M, \theta_M) + T_1(\theta_M, \theta_M) = \theta_M + (2\pi_1 - 1)\theta_M(1 - \theta_M) = 7/12 > W_2$.

Notice that in the proof of Proposition 1 we showed that a necessary condition for the existence of the non-monotonic equilibrium, inequality (15), implies $\mu_{HL} > \theta_M$, that is, the average productivity of the extreme types (the ones choosing the less-informative action) exceeds the expected productivity of the intermediate type (the one choosing the more-informative action).

4.2 Continuous type distributions

Any example with a three-type discrete distribution that satisfies the conditions of Proposition 1 (e.g., Example 1) can easily be transformed into an example with a continuous type distribution where, in a separating equilibrium, intervals of low and high types choose action a_2 and an interval of medium types play a_1.[14]

We now show that similar tri-partite equilibria also exist for a nontrivial set of parameter values for *any* continuous type-distribution. A tri-partite equilibrium is characterized by two cutoffs, A and B with $0 < A < B < 1$, such that the player chooses a_1 if $\theta \in [A, B]$, and plays a_2 otherwise. Denoting the average θ in $[A, B]$ by x_1 and outside $[A, B]$ by x_2, a necessary condition for a tri-partite equilibrium is

$$W_1(x_1, x_1) + T_1(x_1, K) \geq W_2(x_1, x_2) + T_2(x_1, K). \tag{17}$$

That is, the agent with type $\theta = x_1$ prefers action a_1 over a_2—this must be so because $x_1 \in [A, B]$. In what follows, denote the unconditional expected productivity of the agent by $\mu = E[\theta]$. By the continuity and full support of the distribution of θ, we have $\mu < 1$.

As a first step towards proving the existence of a tri-partite equilibrium, we show that there exist thresholds $\bar{\pi}_1$ and $\bar{\pi}_2 : (1/2, \bar{\pi}_1) \to (1/2, \bar{\pi}_1)$ such that when

[13] The assumption $\pi_2 = 1/2$ is a useful simplification for the purpose of calculating the example, but it is certainly not implied by the conditions of Proposition 3.

[14] Details of this construction are available from the authors.

$\pi_1 < \bar{\pi}_1$ and $\pi_2 < \bar{\pi}_2(\pi_1)$, there exists $x_1^* \in (0, \mu)$ such that inequality (17) is satisfied with $x_1 = K = x_1^*$ and $x_2 = \mu$. (The threshold for π_2 must obviously depend on π_1 as $\pi_2 < \pi_1$ by assumption.) Note that in the extreme case, if both actions were uninformative ($\pi_1 = \pi_2 = 1/2$), then the difference between first-period wages, $W_1 - W_2$, and the difference between second-period option values, $T_1 - T_2$, would *both* equal zero. What the following lemma says is that if at least one action is informative, but neither action is too informative, and x_1 is chosen appropriately, then for type $\theta = x_1$ the first-period payoff difference is smaller than the (positive) second-period payoff difference, so this type prefers the more-informative action.

Lemma 2 *There exists $\bar{\pi}_1 \in (1/2, 1)$ and $\bar{\pi}_2 : (1/2, \bar{\pi}_1) \to (1/2, \bar{\pi}_1)$ such that for all $\pi_1 < \bar{\pi}_1$ and $\pi_2 < \bar{\pi}_2(\pi_1) \leq \pi_1$,*

$$W_1(\theta, x_1^*) \leq W_2(\theta, \mu) \text{ for all } \theta \in [0,1], \text{ and} \tag{18}$$

$$W_2(x_1^*, \mu) < W_1(x_1^*, x_1^*) + T_1(x_1^*, x_1^*) - T_2(x_1^*, x_1^*) < 1 \tag{19}$$

hold for some $x_1^* \in (0, \mu)$.

Proof. See the Appendix. ∎

The following proposition establishes the existence of a tri-partite equilibrium in the case of continuously distributed types.

Proposition 2 *Assume that θ has a continuous distribution with full support on $[0, 1]$. There exists $\bar{\pi}_1 \in (1/2, 1)$ and $\bar{\pi}_2 : (1/2, \bar{\pi}_1] \to (1/2, \bar{\pi}_1]$ such that for all $\pi_1 < \bar{\pi}_1$ and $\pi_2 < \bar{\pi}_2(\pi_1) \leq \pi_1$, there exists $K \in (0, 1)$ such that in a separating equilibrium, types $\theta \in [0, A) \cup (B, 1]$ choose a_1 and types $\theta \in [A, B]$ choose a_2, where $0 < A < K < B < 1$.*

Proof. In the proof fix π_1, π_2 and x_1^* such that (18) and (19) hold.

Define

$$\bar{x}_2 = \max \left\{ x_2 | W_2(x_1^*, x_2) + T_2(x_1^*, x_1^*) \leq W_1(x_1^*, x_1^*) + T_1(x_1^*, x_1^*) \right\}.$$

By inequality (19) and monotonicity of W_2 we have $\bar{x}_2 \in (\mu, 1)$.

Define, for all $x_2 \in [\mu, \bar{x}_2]$,

$$C(x_2) = \left\{ K \mid \text{for } \theta = K, \; W_2(\theta, x_2) + T_2(\theta, K) \leq W_1(\theta, x_1^*) + T_1(\theta, K) \right\}.$$

This is the set of outside option levels (K's) such that type $\theta = K$ weakly prefers a_1 to a_2 given the market's beliefs x_1^* and x_2. It is easy to see that $C(x_2)$ is always an interval, $[\underline{c}(x_2), \bar{c}(x_2)]$, that contains x_1^*. Moreover, $\underline{c}(\mu) < x_1^* < \bar{c}(\mu)$ and $C(\bar{x}_2)$ is either $[x_1^*, \bar{c}(\bar{x}_2)]$ or $[\underline{c}(\bar{x}_2), x_1^*]$.

Define, for all $x_2 \in [\mu, \bar{x}_2]$ and $K \in C(x_2)$,

$$D(x_2, K) = \{\theta \mid W_2(\theta, x_2) + T_2(\theta, K) \leq W_1(\theta, x_1^*) + T_1(\theta, K)\}. \tag{20}$$

This is the set of types (θ's) that weakly prefer action a_1 over action a_2 given the market's beliefs, $x_1 = x_1^*$ and x_2, and the outside option, K. Clearly, $D(x_2, K)$ is a non-empty inteval for all (x_2, K) in the domain, and both endpoints of this interval are continuous functions of x_2 and K.

It is easy to see that if K equals either $\underline{c}(x_2)$ or $\bar{c}(x_2)$ then $D(x_2, K) = \{K\}$. Therefore

$$E[\theta \mid \theta \in D(x_2, \underline{c}(x_2))] < x_1^* < E[\theta \mid \theta \in D(x_2, \bar{c}(x_2))].$$

Since the endpoints of $D(x_2, K)$ and the distribution of θ are continuous, the Intermediate Value Theorem implies that there exists $K = K(x_2)$ in the interior of $C(x_2)$ such that

$$E[\theta \mid \theta \in D(x_2, K(x_2))] = x_1^*.$$

$D(x_2, K(x_2))$, which always contains x_1^*, is a non-degenerate interval for all $x_2 \in [\mu, \bar{x}_2)$; however, $D(\bar{x}_2, K(\bar{x}_2)) = \{x_1^*\}$.

Define, for all $x_2 \in [\mu, \bar{x}_2]$,

$$\hat{x}_2(x_2) = \min\{\bar{x}_2, E[\theta \mid \theta \notin D(x_2, K(x_2))]\}.$$

This is a continuous function because D and the distribution of θ are both continuous. Notice that for $x_2 = \mu$, $\hat{x}_2(x_2) = \hat{x}_2(\mu) \in (\mu, \bar{x}_2]$ because $D(\mu, K(\mu))$ is a non-degenerate interval of θ with a conditional expectation $x_1^* < \mu$, while $E[\theta] = \mu$. For $x_2 = \bar{x}_2$, we have $\hat{x}_2(x_2) = \hat{x}_2(\bar{x}_2) = \mu$ because $D(\bar{x}_2, K(\bar{x}_2)) = \{x_1^*\}$ and the distribution of θ is continuous.

Since $\hat{x}_2(x_2)$ is continuous on $[\mu, \bar{x}_2]$ and $\hat{x}_2(\mu) > \mu = \hat{x}_2(\bar{x}_2)$, the Intermediate Value Theorem implies that there exists $x_2^* \in (\mu, \bar{x}_2)$ such that $\hat{x}_2(x_2^*) = x_2^*$.

Finally, we claim that for π_1 and π_2 fixed above and $K = K(x_2^*)$, there exists a separating equilibrium where types $\theta \in [A, B] \equiv D(x_2^*, K(x_2^*))$ choose action a_1 and all other types choose a_2. This is easy to check. The average type choosing a_1 is indeed $E[\theta \mid \theta \in D(x_2^*, K(x_2^*))] = x_1^*$, and the average type choosing a_2

is $E[\theta \mid \theta \notin D(x_2^*, K(x_2^*))] = x_2^*$. Given these market beliefs, the set of types that prefer a_1 over a_2 is exactly $D(x_2^*, K(x_2^*))$ by equation (20). Since x_2^* is in the interior of $[\mu, \bar{x}_2]$ the interval $D(x_2^*, K(x_2^*)) \subset [0,1]$ is non-degenerate and it contains both K and x_1^*. ■

The intuition of how tripartite equilibria are sustained is the following. The low and high types are not keen on learning more about their true productivity, therefore they choose the action that is less informative, but is perceived better by the market in the first period. On the other hand, medium types are interested in updating their beliefs about their productivity, and are willing to be perceived as on average lower types by choosing a more informative action in the first period. For them, this action increases the value of the option to stay or quit in the second period so much so that it outweighs the "stigma" associated with its choice.

The construction of a tri-partite equilibrium can also be illustrated by Figure 3 (ignore the mass-points θ_L, θ_M and θ_H in the picture) . The thresholds $A = \theta'$ and $B = \theta''$ are determined endogenously; types between θ' and θ'' choose action a_1, the more-informative action, while the rest choose a_2.

Example 2 *Suppose θ is uniform on $[0,1]$, $K = 5/11$, $\pi_1 = 5/8$ and $\pi_2 = 1/2$. Numerical calculations reveal that there exists a tri-partite equilibrium such that the player chooses a_1 if and only if $\theta \in [A, B]$, such that $A \approx 0.4224$, $B \approx 0.4976$.*

5 The Structure of Signaling Equilibria

The tripartite equilibria described in the last section have two note-worthy properties: the agent's strategy is non-monotonic, and the average productivity of the agent choosing a less informative action is higher. In this section, we shall show that both features are common to *all equilibria* where both actions are chosen with positive probability. At the end of the section we examine the robustness of the results by briefly discussing certain variants of the model.

5.1 Eagerness to learn indicates self-doubt

First, we show that in any informative equilibrium, action a_2 is associated with on average higher types of the agent. That is, the agent choosing to generate a less precise signal indicates that his expected productivity is higher.

Proposition 3 *In any equilibrium where both actions are played with positive probability we have $x_1 \leq x_2$. That is, a relatively less informative public signal is chosen, on average, by higher types of the player.*

Proof. Suppose towards contradiction that $x_2 < x_1$.

If $x_2 = 0$ then type $\theta = 0$ must be choosing a_2 in the equilibrium. On the other hand, by Lemma 1, $W_2(\theta, x_2) = 0 < W_1(\theta, x_1)$ for all $\theta \in [0, 1]$. Together with $T_2(\theta, K) \leq T_1(\theta, K)$ (which follows from the fact that action a_1 is more informative than a_2) this implies $W_2(\theta, x_2) + T_2(\theta, K) < W_1(\theta, x_1) + T_1(\theta, K)$ for all $\theta \in [0, 1]$. Therefore, type $\theta = 0$ prefers to choose a_1, contradiction.

In the rest of the proof assume $x_2 > 0$.

By (iii) in Lemma 1, $W_2(x_2, x_2) = W_1(x_2, x_2) = x_2$, and by (ii) in Lemma 1, $W_1(x_2, x_2) < W_1(x_2, x_1)$ because $x_1 > x_2$. Similarly, $W_2(x_1, x_2) < W_2(x_1, x_1) = W_1(x_1, x_1) = x_1$. Therefore, for $\theta \in \{x_2, x_1\}$,

$$W_2(\theta, x_2) < W_1(\theta, x_1). \tag{21}$$

Recall that by equation (6) the expected first-period wage, $W_i(\theta, x_i)$, is affine in θ. Therefore, (21) must also hold for all $\theta \in [x_2, x_1]$. Since $T_2(\theta, K) \leq T_1(\theta, K)$, we conclude that for all $\theta \in [x_2, x_1]$,

$$W_2(\theta, x_2) + T_2(\theta, K) < W_1(\theta, x_1) + T_1(\theta, K). \tag{22}$$

Since $W_i(\theta, x_i)$ is affine in θ, either $W_1(\cdot, x_1)$ is steeper than $W_2(\cdot, x_2)$, or $W_2(\cdot, x_2)$ is steeper than $W_1(\cdot, x_1)$. We consider the two cases in turn.

Case 1. Suppose that $W_1(\cdot, x_1)$ is steeper than $W_2(\cdot, x_2)$ is. Then, inequality (22) holds for all $\theta > x_1$ as well. This implies that all types $\theta \geq x_2$ strictly prefer action a_1 over a_2. Hence,

$$\Pr\{\omega = H \mid a(\cdot), a_2\} < x_2,$$

which contradicts condition (3) in the definition of a separating equilibrium.

Case 2. Suppose that $W_2(\cdot, x_2)$ is steeper than $W_1(\cdot, x_1)$. Then, inequality (22) holds for all $\theta < x_2$ as well. This implies that all types $\theta \leq x_1$ strictly prefer action a_1 over a_2. Hence,

$$\Pr\{\omega = H \mid a(\cdot), a_2\} \geq x_1 > x_2,$$

which contradicts condition (3) in the definition of a separating equilibrium. This completes the proof. ■

Proposition 3 rules out the possibility of an equilibrium where a more precise signal-generating action is chosen by on average higher types. The reason why this result may not be obvious is that all else equal, higher types would benefit more from issuing a more informative signal. That is, an agent who is more confident in his productivity is less afraid of the market learning the truth. Moreover, the sensitivity of the agent's gross payoff to his own type depends on the market's beliefs about the productivity of the agent that takes the particular action, and the beliefs are endogenous.

5.2 Non-monotonicity of informative equilibria

The following proposition states that the types of the agent that choose a less precise signal do not dominate the types choosing a more informative one. In other words, the separating equilibrium is *not* a threshold equilibrium.

Proposition 4 *In any equilibrium where both actions are played with positive probability, there exist types $\theta < \theta' < \theta''$ such that both θ and θ'' choose action a_2 while θ' chooses a_1.*

Proof. By equations (4)-(5) and (6), the period 1 expected wage of the agent with type $\theta = 0$ choosing action a_i is

$$W_i(0, x_i) = \frac{(1-\pi_i)\pi_i x_i}{\pi_i x_i + (1-\pi_i)(1-x_i)} + \frac{\pi_i(1-\pi_i)x_i}{(1-\pi_i)x_i + \pi_i(1-x_i)}.$$

By part (ii) of Lemma 1, this expression is increasing in x_i, and by part (iv) of Lemma 1, it is strictly decreasing in π_i as long as $\pi_i > 1/2$. Therefore, in a separating equilibrium, by $\pi_1 > \pi_2$ and $x_1 \leq x_2$, we have $W_1(0, x_1) < W_2(0, x_2)$. The second-period option value for type $\theta = 0$ is zero, therefore that type's total expected payoff from playing action a_2 exceeds the payoff from playing a_1.

This establishes that some type θ sufficiently close to zero strictly prefers, and therefore chooses, action a_2. Since the average type choosing a_2 exceeds the average type choosing a_1, that is, $x_2 > x_1$, it must be the case that some type θ' below x_2 chooses a_1 and some type θ'' above x_2 chooses a_2. This completes the proof. ■

In the proof of Proposition 4 we showed that if the agent is nearly sure that he is not talented (θ is close to zero) then he prefers to send the least informative signal that is associated with on average the highest types. The reason for this is that the lowest type of the agent does not gain from learning about his true productivity—he knows it is low anyway—therefore he might as well choose the least informative action that is rewarded with the highest wage in the short run; moreover, type $\theta = 0$ also likes the fact that a less informative signal is less likely to generate a (correct) low signal about his productivity. The same calculation does not apply to the highest type, $\theta = 1$. Although the agent who is sure that his productivity is high does not gain from learning and likes to be perceived as a higher type, he may prefer a more informative signal that is more likely to generate a (correct) high signal about his talent. Formally, the expected first-period wage of type $\theta = 1$, $W_i(1, x_i)$, is increasing in x_i by Lemma 1, part (ii), but is also increasing in π_i by Lemma 1, part (iv), hence $W_1(1, x_1) < W_2(1, x_2)$ cannot be assured.

5.3 Discussion

In this section we briefly discuss some variants of our base model.

First, one may consider an alternative model where the signal generated by the agent's action in period 1 is observable only to the agent himself. The market still observes his action and pays him a credence wage in the first period; then, in the second period, the agent faces the same in/out decision as he does in the original model. Although this alternative model may not correspond to any real-life situations,[15] all of our results continue to hold there. Intuitively, in this variant, high types have even fewer reasons to choose a more informative action because such an action can no longer signal that the agent is "not afraid of the truth". In any separating equilibrium, choosing a less informative signal displays strength, and only intermediate (on average, lower) types take a more informative action.

Second, one may consider a variant of the model where the agent does not face a second-period decision; instead, there is an exogenous cost of taking action a_2. (Recall that in our original model the "cost" of a_2 arises from a lower second-period payoff.) However, our Proposition 3 continues to hold: the high types' incentive to signal that they are not afraid of the truth is overwhelmed by their incentive to show strength by picking a costlier action. In fact, this result is

[15] For example, it is difficult to imagine that a young artist can create her first work, get rewarded based on whether it is conventional or experimental, and find out its quality without actually showing the piece to the outside world.

so robust that it carries over to models where the two actions are not rankable according to their informativeness.[16]

Finally, let us discuss some of the simplifying assumptions imposed on the model. We limited the agent to choose between two actions, with each action generating a binary signal. These assumptions can be relaxed without compromising any of our results. Perhaps the only technical assumption that is important for our analysis is that the underlying talent of the agent (the state of nature, ω) is also binary. We made this assumption in order to ensure that the agent's private information (the posterior distribution over ω) is one-dimensional. Our analysis could be easily replicated in a model where the agent's type, θ, is a one-dimensional variable indexing a convex set of probability vectors over the values of ω.

6 Conclusions

We have analyzed a new type of signaling game in which a privately, but imperfectly informed sender chooses among observable actions that generate verifiable, stochastic public signals about the payoff-relevant state of nature (e.g., the sender's productivity). The receiver (the "market") uses these signals to infer the state of nature and to pay the sender a corresponding rational-expectations wage. The public signals are also valuable for the agent because learning about the state helps his own decision-making (e.g., concerning his long-run career choice). The result is an interesting trade-off. The agent who is confident about his productivity can provide *stronger evidence* by choosing a more informative signal, but the same action also signals a *lack of self-confidence* because foregoing learning is especially costly for those with self-doubt. These conflicting effects make our game a non-monotonic signaling game where the usual solution methods and results do not apply.

We have shown that in this game there exists a tri-partite equilibrium in which low and high types pool on a less-informative action while intermediate types choose to send a more-informative public signal. The equilibrium action is *non-monotonic* in the agent's type, and the average productivity of an agent choosing a *less informative* action is *greater*. The model sheds light on how signaling and countersignaling can arise and coexist in talent markets. We also believe that our

[16] Details of these arguments are available from the authors.

relatively simple, yet non-monotonic signaling game can be usefully embedded in other applications as well (e.g., in the theory of organizations), and provide a richer framework than traditional models of communication and signaling do.

7 Appendix: Omitted Proofs

Proof of Lemma 1. . (i) If $x_i = 0$ then $w_i^H = w_i^L = 0$ by (4)-(5), and so $W_i(\theta, 0) = 0$ by (6). If $x_i = 1$ then $w_i^H = w_i^L = 1$, hence $W_i(\theta, 1) = 1$ as well.

(ii) From (6),

$$\frac{\partial W_i(\theta, x_i)}{\partial \theta} = (2\pi_i - 1)(w_i^H - w_i^L),$$

which is positive because both terms in the product are positive for $\pi_i > 1/2$.

To see that W_i is strictly increasing in x_i, note that both w_i^H and w_i^L are strictly increasing in x_i provided $\pi_i > 1/2$, and that W_i is just a weighted average of w_i^H and w_i^L.

(iii) For $\theta = x_i$, (7) simplifies to $W_i(x_i, x_i) = \pi_i x_i + (1 - \pi_i)x_i = x_i$.

(iv) Differentiating (7) with respect to π_i yields

$$\frac{\partial}{\partial \pi_i} W_i(\theta, x_i) = \frac{(2\pi_i - 1)(\theta - x_i)(1 - x_i)x_i}{[\pi_i x_i + (1 - \pi_i)(1 - x_i)]^2 [(1 - \pi_i)x_i + \pi_i(1 - x_i)]^2}.$$

The sign of the right-hand side is the same as the sign of $(\theta - x_i)$ because the other terms are all positive. ■

Proof of Lemma 2. Suppose first that $\pi_2 = 1/2$. In this case, $\theta_2^L(\theta) = \theta = \theta_2^H(\theta)$, and so $T_2(\theta, K) = (\theta - K)\,\mathbf{1}_{\theta \geq K}$. Since action a_1 is informative ($\pi_1 > 1/2$), the second-period benefit advantage of action a_1 over a_2, $T_1(\theta, K) - T_2(\theta, K)$, is positive whenever $\theta_1^L(\theta) < K < \theta_1^H(\theta)$, and zero otherwise.

Under the assumption that $\pi_2 = 1/2$, the first-period payoff from choosing action a_2 associated with average type $x_2 = \mu$ is $W_2(\theta, \mu) \equiv \mu$. In order to find x_1^* such that (18) holds, it is sufficient to find x_1^* such that

$$W_1(1, x_1^*) = W_2(1, \mu) \equiv \mu. \tag{23}$$

But for any $\pi_1 > 1/2$ there exists $x_1^* \in (0, \mu)$ satisfying (23) because $W_1(\theta, x_1)$ is continuous in x_1, and by Lemma 1,

$$W_1(1, 0) = 0 < \mu = W_1(\mu, \mu) < W_1(1, \mu).$$

$W_1(\theta, x_1^*)$ is positive and increasing in θ, therefore

$$\mu - W_1(x_1^*, x_1^*) < W_1(1, x_1^*) - W_1(0, x_1^*)$$
$$= \frac{(2\pi_1 - 1)\,\pi_1 x_1^*}{\pi_1 x_1^* + (1-\pi_1)(1-x_1^*)} - \frac{(2\pi_1 - 1)(1-\pi_1)x_1^*}{(1-\pi_1)x_1^* + \pi_1(1-x_1^*)}.$$

By equation (10), $T_1(\theta, K) - T_2(\theta, K)$ peaks at $\theta = K$, where

$$T_1(K, K) - T_2(K, K) = (2\pi_1 - 1)\,K\,(1 - K)\,.$$

Therefore, a sufficient condition for $\mu + T_2(x_1^*, x_1^*) < W_1(x_1^*, x_1^*) + T_1(x_1^*, x_1^*)$ is

$$\frac{(2\pi_1 - 1)\,\pi_1 x_1^*}{\pi_1 x_1^* + (1-\pi_1)(1-x_1^*)} - \frac{(2\pi_1 - 1)(1-\pi_1)x_1^*}{(1-\pi_1)x_1^* + \pi_1(1-x_1^*)} < (2\pi_1 - 1)\,x_1^*\,(1 - x_1^*)\,.$$

Since $\pi_1 > 1/2$, we may cross-divide by $(2\pi_1 - 1)x_1^* > 0$. However,

$$\frac{\pi_1}{\pi_1 x_1^* + (1-\pi_1)(1-x_1^*)} - \frac{1-\pi_1}{(1-\pi_1)x_1^* + \pi_1(1-x_1^*)} < 1 - x_1^*,$$

which holds for any $x_1^* \in (0,1)$ if π_1 is sufficiently close to $1/2$ because the left-hand side tends to zero as π_1 tends to $1/2$. Therefore for π_1 sufficiently close to $1/2$ and $x_1^*(\pi_1)$ satisfying (23), the first inequality in (19) holds. The second inequality in (19) also holds for π_1 close to $1/2$ because $x_1^*(\pi_1) < \mu < 1$ and $\lim_{\pi_1 \to 1/2}\left[T_1(\theta, K) - T_2(\theta, K)\right] = 0$ for all (θ, K).

We established that if $\pi_2 = 1/2$, then there is a threshold $\bar{\pi}_1$ such that for all $\pi_1 \in (1/2, \bar{\pi}_1)$, there exists $x_1^* \in (0, \mu)$ such that conditions (18) and (19) simultaneously hold. By the continuity of all functions involved, the same conclusion holds for all $\pi_1 \in (1/2, \bar{\pi}_1)$ and $\pi_2 \in [1/2, \bar{\pi}_2(\pi_1))$, for some $\bar{\pi}_2(\pi_1) < \pi_1$. ■

References

[1] Avery, Christopher N., and Judith A. Chevalier, "Herding over the career," *Economics Letters*, 63 (1999), 327-333.

[2] Baliga, S., and T. Sjöström, "Arms Races and Negotiations," *Review of Economic Studies*, 71 (2004), 351-369.

[3] Brandenburger, A., and B. Polak, "When Managers Cover Their Posteriors: Making the Decisions the Market Wants to See" *RAND Journal of Economics*, 27:3 (1996), 523-541.

[4] Fang, Hanming, "Social Culture and Economic Performance," *American Economic Review*, 91 (2001), 924-937.

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

[6] Grossman, Sanford (1981): "The Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24, 461-483.

[7] Holmström, Bengt, "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66 (1999), 169-182.

[8] Hörner, J., and N. Sahuguet, "Costly Signaling in Auctions," *Review of Economic Studies*, forthcoming (2007).

[9] Hvide, H. K., "Education and the Allocation of Talent," *Journal of Labor Economics*, 21 (2003), 945-970.

[10] Milgrom, Paul (1981): "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics*, 21, 380-391.

[11] Morris, Stephen, "Political Correctness," *Journal of Political Economy*, 109 (2001), 231-265.

[12] Nelson, P., "Advertising as Information," *Journal of Political Economy*, 82 (1974), 729-754.

[13] Ottaviani, Marco, and Peter Norman Sørensen, "Reputational Cheap Talk," *RAND Journal of Economics*, 37:1 (2006) 155-175.

[14] Ottaviani, Marco, and Peter Norman Sørensen, "Professional Advice," *Journal of Economic Theory*, 126 (2006), 120-142.

[15] Prendergast, Canice, and Lars Stole, "Impetuous Youngsters and Jaded Old-timers," *Journal of Political Economy*, 104 (1996), 1105-1134.

[16] Ross, S. A., "The Determination of Financial Structure: The Incentive-Signaling Approach," *Bell Journal of Economics*, 8 (1977), 23-40.

[17] Scharfstein, David, and Jeremy Stein, "Herd Behavior and Investment," *American Economic Review*, 80 (1990), 465-479.

[18] Spence, Michael, "Job Market Signaling," *Quarterly Journal of Economics*, 87:3 (1973), 355-74.

[19] Teoh, S. H., and C. Y. Hwang, "Nondisclosure and Adverse Disclosure as Signals of Firm Value," *Review of Financial Studies*, 4 (1991), 283-313.

[20] Veblen, T., *The Theory of the Leisure Class.* Macmillan, New York, 1899.

[21] Zahavi, A., "Mate Selection—A Selection for a Handicap," *Journal of Theoretical Biology*, 53 (1975), 205-214.