Will Truth Out?—An Advisor's Quest To Appear Competent *

Nicolas Klein† Tymofiy Mylovanov‡

This version: July 24, 2017

Abstract

We study a dynamic career-concerns environment with an agent who has incentives to appear competent. It is well known that dynamic career concerns create incentives for an agent to be conservative and to tailor his reports towards a commonly held prior opinion. The existing models, however, have focused on short time horizons. We show that, for long time horizons, there exist countervailing incentives for the agent to report his true opinion. In particular, if the agent is sufficiently patient, the time horizon is sufficiently long given the agent's patience, and the quality of the competent expert is high enough given the time horizon and the discount factor, the beneficial long-term incentives overwhelm any harmful myopic ones, and the incentive problem vanishes.

Keywords: Reputational cheap talk, career concerns, advisors, strategic information transmission.

JEL Classification Numbers: C73, D83.

*We are grateful to Editor Prof. Atsushi Kajii and three anonymous referees for their many insightful comments and suggestions, which have greatly improved the paper. We are also thankful to Alessandro Bonatti, Matthias Fahn, Johannes Hörner, Navin Kartik, George Mailath, John Morgan, Andrew Postlewaite, Sven Rady, Larry Samuelson, Joel Sobel, and Satoru Takahashi for very useful comments. Financial support from the National Research Fund, Luxembourg, the German Research Fund (through SFB TR 15), the Fonds de recherche du Québec Société et Culture, and the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. The research reported in this paper is not a result of any for-pay consulting relationship.

†Université de Montréal and CIREQ. Mailing address: Université de Montréal, Département de Sciences Économiques, C.P. 6128 succursale Centre-ville; Montréal, H3C 3J7, Canada, kleinnic@yahoo.com.

‡Department of Economics, University of Pittsburgh; mylovanov@gmail.com.

1 Introduction

Are market forces alone sufficient to solve incentive problems? Fama (1980) argues so based on an agent's reputational concerns, which are thought to be able to discipline his actions. Holmström (1999), by contrast, formalizes the idea that these very reputational concerns may well induce an agent to behave in a way that is detrimental to the principal's interests. In Holmström (1999, Section 3.2), an agent privately observes the success probability of a project, but is reluctant to adopt it efficiently since the market will measure his performance on the marginal adopted project as if it was the average adopted project, much as in Akerlof (1970). As we are interested in the question of whether explicit incentive contracts can be dispensed with, contingent contracts and other sophisticated mechanisms are ruled out and the agent is paid in advance for his perceived competence.

This negative effect of career concerns has so far been shown to exist in environments with short time horizons (see, e.g. Scharfstein and Stein (1990), Trueman (1994), Prendergast and Stole (1996), Holmström (1999), Effinger and Polborn (2001), Levy (2004), Prat (2005), and Ottaviani and Sørensen (2006a, 2006b)).[1] In this paper, we shall show by contrast that, if the agent has a longer horizon, there appear beneficial long-term reputational concerns which counteract the harmful myopic ones analyzed e.g. by Holmström (1999). In our main result, we show that, under certain conditions, an incentive problem which is present for a short time horizon may disappear for a long time horizon, lending some support to Fama's (1980) view in the context of long-term relationships. Thus, the incentives of an agent who wants to leave a good impression only once and those of an agent who wants to leave a good impression over the course of many repeated interactions are not aligned, with the latter more in line with the interests of those whom he wants to impress. As our analysis shows, the reason for this is in the agent's *private* belief about his competence. With many future interactions to come, it is no good for the agent to impress the market in the short run if, in the process, he becomes privately very pessimistic about his ability to impress it in the long run. This would suggest that, if one wanted to interpret static models, like e.g. Ottaviani and Sørensen (2006a,b), in which an agent's reputation enters his utility function, as a "short-hand" representation of a longer interaction where a better reputation induces a higher continuation value for the agent, one ought to assume that the agent's utility depended not just on the *public* belief about his competence but on his *private* belief also.

To be sure, our contribution should not be seen in the (arguably trivial) point that, under certain circumstances, an incentive problem may not arise. Rather, we suggest it be seen in the illustration of the existence of beneficial long-term reputational concerns counteracting the harmful myopic ones, as well as of the importance of a career-concerned agent's *private* belief about his competence, over and above the public belief.

Specifically, in a simple dynamic model featuring a binary policy choice, an iid state of the

[1] Negative reputational effects due to preference uncertainty appear in Morris (2001) and Ely and Välimäki (2003). As the agent knows his preferences in these models, the incentive problem has quite a different structure. See also Morgan and Stocken (2003), Dasgupta and Prat (2006, 2008), Levy (2007), Bourjade and Jullien (2011).

world, and symmetric uncertainty about the agent's binary competence, we demonstrate that there exist countervailing long-term incentives for the agent to be truthful (see Theorem 2). In particular, if the quality of the competent agent is high, incentives to report his information truthfully are fully restored as the number of periods of interaction grows sufficiently large and the agent is sufficiently patient. The key is that the agent can get a really big payoff only by being right many times, something that he has a realistic chance of achieving only if he is of the good type, and then only by reporting his signal truthfully in every period. Thus, to get a very high payoff, it is necessary but not sufficient for the agent to be thought of highly by the market. Indeed, a good reputation today gives him a good wage today. Yet, if he privately knows that he is incompetent, he knows that he will not be able to maintain this good reputation, and the good wages that go with it, for very long. With a long horizon, there thus arises a complementarity between the public belief and the private belief about the agent's competence: Only the histories in which *both* are optimistic are associated with high expected payoffs. Thus, in our model, if the number of periods and the discount factor are both large enough, even an initially pessimistic agent would prefer to bet on the remote prospect that he might be the good type rather than lie, even though lying may seem expedient in the short run. Indeed, the histories in which both beliefs align are accessible only if he conforms to the market's expectation and tells the truth; by lying, he would essentially forgo his, albeit possibly slim, chance of collecting the big payoff associated with being thought of highly in the long run. Hence, in contrast to the existing results for short time horizons, long-run dynamic career concerns could impose sufficient discipline on the agent's behavior.[2]

In our model, the agent is impartial about the policy choices. The good forward-looking incentives generated by career concerns could be used to discipline the agent's behavior in the early periods even if the agent had a conflict of preferences with the principals over the optimal policy. The observation that uncertainty about the agent's competence may alleviate the distortions in communication due to the conflict of preferences is also made in Pavesi and Scotti (2014).

In a somewhat related model, Jullien and Park (2014) show that an infinitely-lived agent can be motivated to tell the truth to convince the market of his high ability, thus eliciting higher future prices, and that uncertainty about the agent's ability facilitates communication despite a conflict of preferences. Jullien and Park's (2014) agent perfectly knows his own ability, while we show that truth-telling may prevail even if the agent has a finite horizon and does not initially have any superior information about his own type, which he is learning over time at an opportunity cost. Furthermore, when our agent is not expected to transmit any information, the market's belief about his ability remains unchanged; Jullien and Park's (2014) agent, by contrast, can never prevent the market from updating its belief about his type.

We motivated our paper with an example from Holmström's (1999) seminal contribution on

[2] This would suggest that in environments where an agent's expertise was essential, arbitrary limits on the duration of relationships might do more harm than good. This observation might be relevant to the ongoing debate about term limits. See Leaver (2009), who demonstrates that longer time horizons can improve the quality of outcomes.

career concerns. Most of the subsequent related literature[3] has, however, followed Holmström's (1999) Section 2 and has focused on environments in which the agent's career concerns reveal themselves through his choice of effort rather than the cheap-talk relay of the agent's private signal.[4] In our environment, by contrast, uncertainty about the agent's competence creates incentives for the agent to lie about his private information (rather than to exert effort), and may slow down learning about the agent's competence.

The rest of the paper is set up as follows. Section 2 presents the model, and introduces necessary notation; in Section 3, we present our results; Section 4 concludes. Proofs are provided in Appendix A; Appendix B discusses a numerical example.

2 Model

Players. There is a long-lived agent (he) who lives for $N \geq 2$ discrete time periods, and a sequence of short-lived principals (she), who each live for a single time period. The horizon N is finite.[5]

Actions. In each period $t \in \{1, \cdots, N\}$, the principal first makes a wage offer $v_t \geq 0$ to the agent. In line with the career-concerns literature, the wage level is exogenously given by a competitive labor market. (See below for details.) The agent then decides whether to accept to work for the principal in the current period at the specified wage level ($k_t = 1$) or not ($k_t = 0$). If he accepts the wage offer in period t, the agent then privately observes the realization of a noisy signal $\tilde{s}_t \in \{0, 1\}$ about the current state of the world $\omega_t \in \{0, 1\}$. States of the world are iid across periods; for all t, the probability that $\omega_t = 1$ is $p \in (0, \frac{1}{2})$. After observing the signal, the agent sends a cheap-talk report $\hat{s}_t \in \{0, 1\}$ to the principal. The principal then chooses a policy $y_t \in \{0, 1\}$, after which the state ω_t is publicly observed, and, if $t < N$, play moves to the next period $t + 1$. (If $t = N$, the game then ends.) If the agent is not employed in the current period ($k_t = 0$), the principal immediately chooses a policy $y_t \in \{0, 1\}$, after which the state ω_t is publicly observed, and play moves into the next period (or, if $t = N$, the game ends).

Information. The agent observes the entire past history, i.e. the past wage offers and employment decisions, the signals and his reports in those periods in which he was employed, as well as past policy choices and states of the world. Formally, for $t \geq 2$, an agent-history of length $t - 1$ is the tuple $h_A^{t-1} := \left(\{v_\tau\}_{\tau=1}^{t-1}, \{k_\tau\}_{\tau=1}^{t-1}, \{\tilde{s}_\tau\}_{\tau=1}^{t-1}, \{\hat{s}_\tau\}_{\tau=1}^{t-1}, \{y_\tau\}_{\tau=1}^{t-1}, \{\omega_\tau\}_{\tau=1}^{t-1}\right)$, where we set $\tilde{s}_\tau = \hat{s}_\tau = \emptyset$ in all periods $\tau \leq t - 1$ such that $k_\tau = 0$. Let $\mathcal{H}_A^{t-1}$ be the set of all agent-histories of length $t - 1$. (We set $\mathcal{H}_A^0 = \{\emptyset\}$.) A pure strategy for the agent specifies what wage offers to accept in each period as a function of the previous history, as well as a report in all periods in which the agent is hired, as a function of the previous history, the current wage and

[3]See Bonatti and Hörner (2017) and the references therein.

[4]This is e.g. the case for the recent contributions by Bonatti and Hörner (2017) and Board and Meyer-ter-Vehn (2013), who also analyze the problem of an agent who is concerned about his reputation in the long run.

[5]We briefly discuss the infinite-horizon case at the end of the next section.

the realization of the current signal. Formally, it is a sequence of mappings $\{\sigma_t^A\}_{t=1}^N$, where, for each $t \in \{1, \cdots, N\}$, $\sigma_t^A = (k_t, \hat{s}_t)$, with $k_t : \mathcal{H}_A^{t-1} \times \mathbb{R}_+ \to \{0,1\}$, $(h_A^{t-1}, v_t) \mapsto k_t(h_A^{t-1}, v_t)$ and $\hat{s}_t : \mathcal{H}_A^{t-1} \times \mathbb{R}_+ \times \{0,1\}^2 \to \{0,1,\emptyset\}$, $(h_A^{t-1}, v_t, k_t, \tilde{s}_t) \mapsto \hat{s}_t(h_A^{t-1}, v_t, k_t, \tilde{s}_t)$, with the restriction that $k_t = 0 \Rightarrow \hat{s}_t(h_A^{t-1}, v_t, k_t, \tilde{s}_t) = \emptyset$ and $k_t = 1 \Rightarrow \hat{s}_t(h_A^{t-1}, v_t, k_t, \tilde{s}_t) \in \{0,1\}$. A mixed strategy for the agent is a probability distribution over the set of his pure strategies.

While each principal observes all public information from previous periods, a principal-history is still coarser, as the principal does not observe the signal realizations $\{\tilde{s}_\tau\}_{\tau=1}^{t-1}$. Formally, for $t \geq 2$, a principal-history of length $t-1$ is the tuple $h_P^{t-1} := \left(\{v_\tau\}_{\tau=1}^{t-1}, \{k_\tau\}_{\tau=1}^{t-1}, \{\hat{s}_\tau\}_{\tau=1}^{t-1}, \{y_\tau\}_{\tau=1}^{t-1}, \{\omega_\tau\}_{\tau=1}^{t-1}\right)$, where $\hat{s}_\tau = \emptyset$ in all periods $\tau \leq t-1$ such that $k_\tau = 0$. A strategy for the principal specifies a policy choice as a function of the previous history and, if the agent is hired, his report. Formally, the period t-principal's strategy is a mapping y_t specifying her policy decision as a function of the previous history, the wage offer, the agent's employment decision and his report, i.e. $y_t : \mathcal{H}_P^{t-1} \times \mathbb{R}_+ \times \{0,1\}^2 \to \{0,1\}$, $(h_P^{t-1}, v_t, k_t, \hat{s}_t) \mapsto y_t(h_P^{t-1}, v_t, k_t, \hat{s}_t)$. A mixed strategy for the period t-principal is a probability distribution over the set of her pure strategies.

The precision of the agent's signals $\{\tilde{s}_t\}_{t=1}^N$, which we call the agent's *competence*, can be high or low, and is initially unknown. The high-quality signal is correct with probability $q \in (1-p, 1]$ while the low-quality signal is correct with probability $r \in [1/2, 1-p)$;[6] our main result will require that q be close to 1. These probabilities are time-invariant and commonly known. The signals are iid across periods. The agent's competence is constant over time.

The parties start out with a common prior, assessing the agent to be competent (i.e., his signals to be of high quality) with probability $\alpha \in (0,1)$. As the game proceeds, both the agent and the market update their respective belief about the agent's competence. We denote by α_t the market's, i.e., the principals', belief about the agent's competence at the beginning of period t; we refer to it as the agent's *reputation.* This belief could well differ from the agent's, for the market's information is in general a garbling of the agent's information, whenever the latter does not fully reveal his private signal. In truthful equilibrium, where the agent fully reveals his private information, this issue arises only off the path of play. Indeed, the filtration generated by the principal-histories is coarser than that generated by the agent-histories, as the agent has the benefit of additionally knowing the signals he has observed.

Payoffs. Each principal's period payoff is equal to u, which we normalize to 1, if the policy matches the state and 0 otherwise, minus the wages paid to the agent. In line with the career-concerns literature, we assume that the agent is paid upfront at the beginning of each period and that the labor market is competitive. Thus, the agent's wages are such that, in any period he is employed, the principal's expected profit (with respect to the public information)

[6]That is, for all $t \in \{1, 2, \cdots, N\}$, $\Pr(\tilde{s}_t = 1 | \omega_t = 1) = \Pr(\tilde{s}_t = 0 | \omega_t = 0) = \psi$, where $\psi = q$ ($\psi = r$) if the signal is of high (low) quality. Our assumption that $q > 1-p$ ensures that a competent agent's best estimate of the state is more precise than the principal's; the assumption that $r < 1-p$ ensures that a bad agent would be strictly better off relying on the public information than on his signal. Without the former assumption, the agent would be useless. Without the latter assumption, truth-telling would be a (weakly) dominant strategy for the agent, whatever the belief about his competence.

is the same as if she did not employ him, i.e. $1-p$; the agent reaps the entire surplus he generates. The wages will thus depend on the agent's reputation as well as his equilibrium strategy. In particular, if there is an equilibrium in which the agent will truthfully reveal his signal in every period in which he is employed, the surplus he generates for the period t-principal hiring him, given his reputation is α_t, is either 0 (if the principal optimally ignores his advice) or $\alpha_t q + (1-\alpha_t)r - (1-p)$ (if the principal optimally follows his advice). Thus, in a truthful equilibrium, his wages in period t will be given by $v(\alpha_t) := \max\{0, \alpha_t q + (1-\alpha_t)r - (1-p)\}$ in all periods in which he is employed.[7] In those periods in which he is unemployed, he receives an exogenous and fixed outside option of $v_0 \geq 0$, which does not depend on his competence. We assume that the agent's outside option is not very high, i.e. $v_0 < \alpha p$. The agent discounts future payoffs with the discount factor $\delta \in (0,1]$.

The agent is impartial regarding the policy and his payoff is equal to his discounted expected wage stream,

$$\mathbb{E}\left[\sum_{t=1}^{N} \delta^{t-1}\left((1-k_t)v_0 + k_t v_t\right)\right],$$

where the expectation is with respect to the processes $\{k_t\}$ and $\{v_t\}$.

The period t principal's ex ante expected payoff, given her strategy and the agent's reputation α_t, is given by

$$\mathbb{E}\left[\left((1-k_t)\Pr(y_t=\omega_t) + k_t\left(\Pr(y_t=\omega_t|\hat{s}_t) - v(\alpha_t)\right)\right)\right],$$

where the expectation is with respect to the random variable k_t, and the probabilities that the principal's policy choice matches the state.

Our solution concept is perfect Bayesian equilibrium; there are no long-term contracts or other ways for the principals or the agent to commit to a certain course of action. The players are allowed to use mixed strategies.

3 Results

3.1 The First-Best Benchmark

As our first-best benchmark, we consider a hypothetical environment in which the principals and the agent act cooperatively so as to maximize the sum of their expected discounted payoffs. For $t \geq 2$, a (first-best) history of length $t-1$ is the tuple $h^{t-1} := \left(\{k_\tau\}_{\tau=1}^{t-1}, \{\tilde{s}_\tau\}_{\tau=1}^{t-1}, \{y_\tau\}_{\tau=1}^{t-1}, \{\omega_\tau\}_{\tau=1}^{t-1}\right)$, where we set $\tilde{s}_\tau = \emptyset$ in all periods $\tau \leq t-1$ such that $k_\tau = 0$. Let $\mathcal{H}^{t-1}$ be the set of all first-best

[7] Thus, we assume that the competitive market prices the agent's services taking into account that, in each period, the principal who hires him will use the agent's expertise optimally, by, depending on the agent's reputation, either following his advice or ignoring it. Clearly, in any equilibrium, every principal will behave in this myopically optimal way (and, in particular, will only consider playing a non-degenerate mixed strategy when both pure strategies give her the same payoff).

histories of length $t-1$. (We set $\mathcal{H}^0 = \{\emptyset\}$.) A strategy specifies whether to acquire a signal (hire the agent) or not, as well as the policy choice, as a function of the history. Formally, a strategy in this setting is a sequence of mappings $(\sigma_t)_{t=1}^N$, where, for each $t \in \{1, \cdots, N\}$, $\sigma_t = (k_t, y_t)$, where, in a slight abuse of notation, we write $k_t : \mathcal{H}^{t-1} \to \{0,1\}, \;\; h^{t-1} \mapsto k_t(h^{t-1})$ for the employment decision as a function of the history, and $y_t : \mathcal{H}^{t-1} \times \{0,1\}^2 \to \{0,1\}, \;\; (h^{t-1}, k_t, \tilde{s}_t) \mapsto y_t(h^{t-1}, k_t, \tilde{s}_t)$ for the policy choice as a function of the history as well as the current employment decision and signal realization. Clearly, optimality requires a policy choice of $y_t = 0$ in all periods in which the agent is not employed; when the agent is employed, the optimal policy is that which coincides with the state that seems more likely given the realization of the signal $\tilde{s}_t$. The optimal employment decisions $\{k_t\}_{t=1}^N$ are those which maximize

$$\frac{1-\delta^N}{1-\delta}(1-p) + \mathbb{E}\left[\sum_{t=1}^{N} \delta^{t-1}\left((1-k_t)v_0 + k_t \max\{0, \alpha_t q + (1-\alpha_t) r - (1-p)\}\right)\right],$$

where the expectation is taken with respect to the processes $\{k_t\}_{t=1}^N$ and $\{\alpha_t\}_{t=1}^N$ under the filtration generated by the first-best histories.

Our first-best problem is in fact an experimentation problem with a bandit with a safe arm and a risky arm. Indeed, not hiring the agent gives a safe period payoff of $v_0 + 1 - p$. Hiring the agent can be thought of as pulling the risky arm, the quality of which can be high or low and is initially uncertain. Note that, for low α_t, it may well be optimal to acquire a signal at an opportunity cost of v_0 even if this signal is optimally disregarded for the policy choice in the current period. The reason is that hiring the agent now will provide valuable information concerning his quality, which information may subsequently be parlayed into better policy choices in the future.

The first-best employment decisions have the property that, if the agent is not employed in a given period, he will not be employed in any subsequent period either, since no new information is learnt about the agent in this case (and there is one period less potentially to benefit from the agent's expertise). In any period t, the first-best employment decision weighs the entire discounted expected benefit in the periods $t, t+1, \cdots, N$ of employing the agent in the current period, taking into account the value of the option of terminating him at some future date, against $\frac{1-\delta^{N-t+1}}{1-\delta} v_0$, the payoff from not employing the agent. The following Proposition summarizes the structure of a first-best policy.

Proposition 1 (First-Best Policy) *There exists a first-best policy that is characterized by a sequence of thresholds $(\bar{\alpha}(N-t))_{t=1}^N$ such that the agent is employed in period t if and only if $\alpha_t \geq \bar{\alpha}(N-t)$. The thresholds $\bar{\alpha}(N-t)$ are non-decreasing in t, and constant if $N = \infty$.*

Proof: See Appendix A.

Note that, if $q = 1$, the first-best employment rule has a particularly simple structure: If the agent is employed in period 1, he will continue to be employed as long as his previous signals have all been correct; at his first mistake, he is fired for good. This simple structure continues to apply for q close to 1, which is the case on which we will focus below; more precisely, for any couple (N, δ) of the horizon and the discount factor, there exists a $\bar{q}_{N,\delta} \in (0, 1)$ such that for $q \geq \bar{q}_{N,\delta}$, the first best has this simple structure.

3.2 Incentive Compatibility of the First Best

We now return to the original model, in which the signal is *privately* observed by the agent whenever he is hired. We are interested in sufficient conditions guaranteeing the existence of an equilibrium which coincides with our first-best benchmark. We begin by stating the following

Definition The first-best outcome is *incentive compatible* if there exists a perfect Bayesian equilibrium in which, for every realization of the agent's signals and states of the world, the employment decision and the policy choice coincide with that of the first-best outcome.

Let α_t^* be the reputation of the agent in period t given that he is expected always truthfully to reveal the realization of his signal, and given that he has been correct in all preceding periods,

$$\alpha_t^* := \frac{\alpha q^{t-1}}{\alpha q^{t-1} + (1-\alpha) r^{t-1}}.$$

Then, if his reputation is α_t^*, in a first-best equilibrium, the agent's wages in period t are equal to

$$v_t^* := \max\{0, r + \alpha_t^*(q - r) - (1 - p)\}.$$

As the previous literature has emphasized, career concerns might create incentives for agents to distort their private information. Readers familiar with the logic of this argument can skip to the paragraph preceding Theorem 2. To provide an illustration of the agency problem which can arise in our setting, we provide a simple two-period example. Let us assume $N = 2$, $\delta = 1$, and let

$$r + \alpha(q - r) - (1 - p) < 0 < v_2^*. \tag{1}$$

Note that the first part of Inequality (1) implies both that the first-period principal should ignore the agent's signal when deciding on the policy, and that the agent has a better chance of being correct in period 1 by sending the message $\hat{s}_1 = 0$ regardless of the signal realization. Furthermore, the inequality implies that the agent's information becomes valuable if and only if his information proves to be correct in the first period. Consequently, if v_0 is sufficiently small, the first-best outcome is for the agent to work for the principal in the first period and work for the principal in the second period if and only if his first-period report is correct. As our discussion below will show, the first best is not incentive compatible in this case. The reason is that it is impossible to induce the agent truthfully to reveal his first-period signal, ruling out sorting between agents based on whether their first-period signal was correct. The presence of agents whose first-period signal was incorrect in turn depresses wages in the second period, possibly to the extent of shutting down the market completely. We summarize this in the following

Result Suppose (1) and $\delta = 1$. If $v_0 < \frac{r+\alpha(q-r)}{1+r+\alpha(q-r)} v_2^*$, the first best is not incentive compatible in a setting with $N = 2$.

By (1), the first-best policy choice in the first period is $y_1 = 0$ regardless of the signal realization. In the second period, the first-best policy will conform to the signal if the agent's signal was correct in the first period and he is consequently employed in the second period; otherwise, it will be $y_2 = 0$. Thus, even though, in the first-best policy, the first-period signal is not used to determine the policy choice in the first period, acquiring the signal is still worth the cost v_0, as doing so yields information concerning the precision of the signal, which enables a better policy choice in the second period. The surplus created by the agent's employment in a first-best world, in which the principals can observe the realization of the agent's signal, thus amounts to

$$w = \underbrace{-v_0}_{\text{value added in } t=1} + \underbrace{(r + \alpha(q-r))}_{\text{prob } \tilde{s}_1 = \omega_1} \times \underbrace{(v_2^* - v_0)}_{\text{value added in } t=2}$$

and is positive if and only if

$$v_0 \leq v := \frac{r + \alpha(q-r)}{1 + r + \alpha(q-r)} v_2^*$$

Now, imagine that the agent observes $\tilde{s}_1 = 1$ in the first period. As we have seen, by (1), the principal's optimal policy in the first period is $y_1 = 0$ independently of the agent's information, implying that the posterior beliefs after signal 1 nevertheless assign a higher probability to state $\omega_1 = 0$. Thus, if the market were to believe the agent's reports, the agent would maximize the probability of getting the positive wage of v_2^* in the second period by lying and reporting $\hat{s}_1 = 0$. As a result, the agent's best response entails a report of 0 in period 1 irrespective of the observed signal. Nevertheless, as there are no incentive problems in the second period, the agent's value for the principal in the second period can be positive since the agent can condition his recommendation on his *private* belief about his competence. This value is maximized if the agent reports his second-period signal truthfully if and only if his first period signal is correct. Yet, as the agent's first-period message is uninformative about his competence, the market does not update his reputation from α regardless of whether his report was correct or not. Thus, the wage offer in the second period will amount to $(r + \alpha(q-r))v_2^*$ (instead of v_2^* in case of a first-period success in the first best), as the market will assess the likelihood of the agent's signal having been correct in the first period as $r + \alpha(q-r)$ regardless of whether the first-period report matched the state or not. Since the agent is employed in both periods, the ex ante expected surplus created by the agent's employment amounts to

$$\hat{w} = \underbrace{-v_0}_{\text{value added in } t=1} + \underbrace{(r + \alpha(q-r))v_2^* - v_0}_{\text{value added in } t=2},$$

and is positive, if and only if

$$v_0 \leq \hat{v} := \frac{r + \alpha(q-r)}{2} v_2^*.$$

It follows that if $v_0 > \hat{v}$, the agent will not accept employment in equilibrium and the market will not be able to benefit from the agent's expertise in the last period. The difference to the first-best policy is that the principal's wage offer in the second period cannot condition on the realization of the agent's signal in the first period. The agent being the residual claimant of the surplus in equilibrium, the inefficiency shows up in a lower second-period wage as compared to the first best ($(r+\alpha(q-r))v_2^*$ instead of v_2^*), thus making the agent's outside option of declining employment in both periods, giving him $2v_0$, seem relatively more attractive. In summary, if $\hat{v} < v_0 < v$, the first-best outcome is not incentive compatible. Furthermore, even if $v_0 \leq \hat{v}$ and the agent is employed for sure in both periods in equilibrium, the employment outcome does not condition on the correctness of the agent's signal in the first period and, hence, the employment decisions do not coincide with the first best. Moreover, the surplus, and hence the agent's wages, are lower.

To summarize our example, the first best has the policy following the signal in the second period if and only if the signal in the first period was correct. In the first period, by contrast, the parties do not yet trust the quality of the signal enough for it to overcome their prior belief; i.e. the optimal policy in the first period is $y_1 = 0$ regardless of the realization of the signal. Therefore, if the principals expect the agent to relay the signal realization truthfully, the agent will claim that the signal realization in the first period corroborates the parties' prior belief in order to maximize his chances of being employed in the second period. Thus, there is no equilibrium in which the agent reports his signal truthfully in the first period and he is employed in the second period if and only if his first-period report matched the state; i.e. the first best is not incentive compatible.

While the analogy is not perfect, the logic of this example is familiar from Holmström (1999), and in particular, that paper's Section 3.2. Indeed, just as Holmström's (1999) agent will always refrain from investing in the first period, thus freezing the public belief at the prior, in our two-period example, the agent will not want to reveal his signal, thus shutting down equilibrium learning about his competence. Of course, while Holmström's (1999) agent could unilaterally shut down learning by not investing, the freezing of the belief is an equilibrium phenomenon in our setting: Indeed, for the agent to be able to freeze the public belief, it is necessary that the market *expect* him not to reveal his signal. Furthermore, while in Holmström's (1999) example, the agency costs arise as a result of forgone investment opportunities in the first period, in our example, they show up as a depressed wage offer in the second period. The root of the inefficiency in both settings is in the agent's incentives to distort the information the market receives in a way that benefits the agent's reputation. Furthermore, in both instances, the problem is due to the assumption that contingent wages are not possible and the agent is paid upfront for his reputation.

In our main result, we show that if the discount factor is sufficiently high, then there exists a minimal length of time, such that for any game longer than this minimal length, there exists a set of information qualities for which the first best is incentive compatible; i.e., the agent becomes willing truthfully to report his signals, and the agency problem disappears. This result qualifies the intuition of career-concerns models based on two-period interactions, which suggest that a career-concerned expert who cares about his reputation *in the next period* makes the report the principal wants to hear. The key insight of this paper is that, if a career-concerned agent cares about appearing competent *in the long run*, these long-run reputational concerns may overwhelm the harmful short-term reputational concerns, and the agency problem may disappear completely.

Theorem 2 (Vanishing Incentive Problem) *For any given p, α, r, and v_0 such that $v_0 < \alpha p$, there exists a $\check{\delta}_0 < 1$ such that, for all $\delta \in (\check{\delta}_0, 1)$, there exists an integer N_δ such that, for all $N > N_\delta$, there exists a $q_{N,\delta} \in (1-p, 1)$ so that the first best is incentive compatible for all $\delta > \check{\delta}_0$, $N > N_\delta$ and $q \geq q_{N,\delta}$.*

Proof: See Appendix A. ∎

To get an intuition for the forces behind this result, note that if, in the first-best outcome, it is optimal to choose the policy that coincides with the signal, reporting his signal truthfully is clearly incentive-compatible for the agent, as it maximizes his chances of being correct. So, suppose, on the contrary, that the first best prescribes a policy choice of $y = 0$ regardless of the signal, and suppose that, in equilibrium, the agent is expected to reveal his signal truthfully. As we have seen above in our discussion of the two-period example, the first best would not be incentive compatible in this case if the game ended in the next period.

But suppose that the agent considers a longer horizon, and, to get an intuition, suppose that $q = 1$, i.e., the competent expert's signal is never wrong. Thus, if the expert's report turns out to be wrong, the market will be certain that he is incompetent, and he will therefore never be able to command positive wages again. If his report is correct, his public reputation increases. If he was honest with his report, his private belief is enhanced as well in this case; but, if he was lying, he *privately* learns that he is incompetent, and that, no matter what he will do in the future, he is bound to make a mistake, and hence to be found out, soon. In other terms, if the agent lies, he is either shut out of the market right away (because his signal was in fact correct and the market now believes him to be incompetent for sure), or he privately learns that he is the bad type, which is bad news about the expected wage offers he will get in the future. If, however, he is telling the truth, the public and private beliefs will always be in sync. If his report is correct, both his reputation and his own assessment of his competence will be enhanced. If his report turns out to be wrong, both he and the market will know that he is the bad type and he will be fired; yet, conditional on his being incompetent, he would have been found out soon anyhow. His only realistic chance of receiving a high payoff is for him to be competent *and* for the market to believe in his competence over the long run. In fact, the longer the horizon the less of a chance the expert has to fool the market until the end, and therefore the relatively more attractive honesty will look to him. Honesty indeed affords the agent a positive chance, namely the initial probability that he is competent α, of being guaranteed the really big payoff associated with being thought of highly until the end; the probability of achieving this feat by lying, in contrast, vanishes as the number of periods becomes large. Thus, while a longer horizon and an increasing discount factor increase the stakes at play for the agent and hence make it all the more important for him to avoid mistakes, it becomes increasingly less likely, as the horizon increases, that the agent can secure these additional payoffs by lying. Therefore, honesty becomes *relatively* more attractive than lying, as the horizon increases. In the words commonly attributed to Abraham Lincoln: "You can fool all the people some of the time, and some of the people all the time, but you cannot fool all the people all the time."[8]

This increase in the importance of the complementarity in private and public beliefs is the key to understanding Theorem 2, which shows that, if the future is important enough (i.e. the horizon N is long enough and the discount factor δ is close enough to 1), then incentives for truth-telling are restored. Of course, our reasoning above relied on the last period being far enough removed. Yet, since the agent becomes more optimistic about the accuracy of his signal with each realization that corresponds to the state, misreporting is profitable for the agent early on if ever. Thus, all that remains to be shown is that *initial* optimism is sufficient for the first best to be incentive-compatible, which, for an agent who is far-sighted enough and has a long enough horizon, is ensured by our assumption $v_0 < \alpha p$. To get an intuition for the role this assumption plays in the proof, note that a good agent will eventually get wages approaching p per period in the long run if he always reports his signal truthfully.[9] The assumption thus guarantees that an agent who hypothetically only cared about his long-run prospects is initially optimistic enough to prefer betting on his ability of eventually securing himself wages of p per period over his outside option v_0. In the proof of Theorem 2, we establish by a continuity argument that a similar logic continues to apply for $q < 1$ but close enough to 1.

As is also easily seen from the proof, if $q = 1$, the result continues to hold in an infinite-horizon model. That is, for any given p, α, r, and v_0 such that $v_0 < \alpha p$, there exists a $\check{\delta} < 1$ such that the first best is incentive compatible in the infinite-horizon game ($N = \infty$) if $q = 1$ and $\delta > \check{\delta}$.

The difficulty with extending the result to $q < 1$ and $N = \infty$ lies in understanding the first-best policy. Indeed, for any integer k, there exists a neighborhood of discount factors close to 1 such that the principals' wage offers would not drop to 0 *forever* after k mistakes. Whatever $q < 1$ may be, there always remains some chance that, even after k mistakes, the agent may turn out to be competent in the distant future. However small this chance may be, there exists a discount factor $\delta_q < 1$ such that, for $\delta > \delta_q$, the agent attaches sufficient weight to this chance in the distant future that he would be willing to continue working even after a k-th mistake. The description of the first-best policy thus becomes much more complicated, and we have been unable to characterize it or to show that it is implementable.

For the case in which the first best has the property that it is optimal never to hire the agent again after a first mistake, the result in Theorem 2 does not depend on our assumption of a common prior about the agent's competence. Indeed, suppose that the agent initially had *some* private information concerning his competence. Our argument establishing the incentive compatibility of truthtelling would go through unchanged if the agent's private prior belief $\breve{\alpha}$ satisfied our assumption $v_0 < \breve{\alpha} p$. Moreover, the market would expect the agent to tell the

[8]See e.g. https://www.brainyquote.com/quotes/quotes/a/abrahamlin110340.html

[9]Indeed, in this case, $\alpha_t^* \to 1$ and $v_t^* \to p$, as $t \to \infty$.

truth on the equilibrium path if it knew that the agent's prior belief satisfied this condition. Therefore, the first-best outcome will continue to be incentive compatible in this case.[10] If the agent were *perfectly* informed about his competence, he would trivially best respond by always reporting a signal realization of 0 if he knew himself to be of the low type, while truthfully reporting his signal if he knew himself to be of the high type.

In the statement of Theorem 2, the thresholds for the number of periods and the precision of the good expert's information are chosen in sequence; i.e., the threshold for the number of periods, N_δ, depends on the actual discount factor δ (rather than just the threshold $\check{\delta}_0$), and the threshold for the precision of the good expert's information $q_{N,\delta}$ depends on N and δ (rather than just on the thresholds $\check{\delta}_0$ and N_δ). If we fix $q = 1$, the result can be strengthened to hold for uniform bounds, as the following proposition shows.

Proposition 3 (Uniform Bounds For $q = 1$) *Let $q = 1$. For any given p, α, r, and v_0 such that $v_0 < \alpha p$, there exists a $\delta_0 < 1$ such that the first best is incentive compatible if $\delta > \delta_0$ and $N > N_{\delta_0}$, for some finite N_{δ_0}.*

PROOF: See Appendix A.

This stronger result is not true in general if $q < 1$. Indeed, consider any set of parameters (q, δ, N), with $q < 1$, such that the first best is incentive compatible and calls for firing the agent after a first mistake. Now, keep q fixed but increase N and/or δ. At some point, the first best will no longer call for firing the agent after a first mistake, because, given the higher N and δ, there is now some tiny probability with which the agent will be able to make up for a first mistake by subsequently being correct for many times in a row, after which his advice may be very valuable. Once the structure of the first best changes in this way, the first best will not generally continue to remain incentive compatible. Indeed, consider a history at which the agent's reputation is so low that the first best calls for firing him if he made a mistake in the current round. In this case, the agent wants to maximize the probability of being correct in the *current* period, and the harmful myopic reputational concerns carry the day over the beneficial long-term reputational concerns. We give a simple numerical example of this effect in Appendix B.

4 Conclusion

We have investigated the dynamic interaction between a sequence of short-lived principals and a long-lived agent of unknown ability who privately observes a potentially decision-relevant signal. As he only cares about his reputation insofar as it translates into higher lifetime wages, the

[10] By contrast, if the agent chose the policy, the market's belief about the agent's prior belief would matter for the agent's wages. Therefore, the agent might have an incentive to manipulate the market's belief about his prior belief by following his signal sooner, thereby signaling a higher prior confidence in his capabilities. A full analysis of this case is outside the scope of this paper.

agent may have incentives strategically to manipulate the cheap-talk relay of his signal to the principals. We have shown that if a competent agent makes mistakes not too frequently, players are patient enough, and the number of periods is large enough, the agent's career concerns no longer give rise to any incentive problems, and the first best becomes implementable.

The agent is impartial about the policy choice in our model and cannot improve the precision of the signal by exerting additional effort. We also assume an information structure that creates incentives for the agent to herd on the prior.[11] We make these assumptions to help render the intuition behind the positive forward-looking incentives created by career concerns as clear as possible.

The positive result in this paper is robust to a number of modifications of the model. It extends to settings in which the agent receives a fixed, possibly intrinsic, benefit from being employed, or the agent rather than the principal chooses the policy. In fact, the idea that the market can 'threaten' the agent with beliefs that attribute any suboptimal policy recommendation to incompetence on the agent's part is quite general. It would e.g. persist in environments in which the agent is partial about the policy, or observes signals generated by asymmetric signal distributions.

Perhaps the most restrictive assumption in our model is the binary structure of the policy, the state, and the agent's type spaces. We believe that the simplicity of the model makes the intuition transparent, that this intuition is robust and transferrable to other environments, and that the positive effect of long-term countervailing incentives will manifest itself in other, richer, models. In particular, the logic of the arguments should extend to environments in which there are multiple discrete policies and levels of competence, provided that the agent is valuable to the market only if his competence is sufficiently high. An interesting question is the nature of the limit result if the agent's competence level has continuous support.

We have left open the question of equilibrium outcomes when the first best is not incentive compatible. While the agent's report is cheap talk, ours is not a typical cheap-talk setting in that the acquisition of a signal, and hence the opportunity to send a message, comes at a cost of v_0 per period. This in particular rules out a *babbling equilibrium*, which always exists in standard cheap-talk settings and which is characterized by the agent's never transmitting any useful information while hired. For the case in which q is close to 1, our two-period example, however, would suggest the existence of an equilibrium with an initial "grace stage" during which the agent is not expected to report his signal truthfully until some period $\hat{\tau} - 1$, after which he is retained if and only if his report is correct in every period, where $\hat{\tau}$ is the period in which the first-best policy starts following the signal given that all previous signals have been correct. Of course, the existence of such an equilibrium would require the initial babbling phase not to be too costly, i.e. it is necessary for v_0 to be low enough. In our two-period example, it requires $v_0 \leq \hat{v}$.

[11] Imagine that an incompetent expert always observes signal 0. In this environment, the agent might have contrarian reporting incentives since, in truthful equilibrium, if it exists, a report of the *a priori* unlikely signal 1 enhances the public belief about the agent's competence.

We have restricted ourselves to a setting with a single task. In the real world, however, one would think that there was some sorting of agents among tasks, with the more competent agents working on the more demanding tasks. This would introduce an additional allocative benefit to finding out the agent's type early on, making lying on the agent's part all the more harmful.

One could think of a number of other possible extensions, such as allowing for the agent to be replaced with a new agent from a pool of experts, or introducing learning by doing on the agent's part. It might also be interesting to introduce multiple, possibly even a continuum of, agents into our setup. We commend these questions to future research.

Appendix

A Proofs

Proof of Proposition 1

Claim 1: There exists a first-best policy such that the agent is never re-hired once he has been fired.[12]

Let us first assume that the horizon $N < \infty$. We prove the statement by induction over $k := N - t$. The statement is trivially true for $k = 0$. Now, let us posit the induction hypothesis that it is true for stage $k - 1$. The induction hypothesis immediately gives us our conclusion unless an optimal strategy is given by the strategy $\hat{\sigma}$ according to which the agent is not hired in period $N - k$, but hired again in period $N - k + 1$, to be fired forever after some (random) period $\hat{t} \in \{N - k + 1, N - k + 2, \cdots, N\}$ (if $\hat{t} = N$, the agent is never fired). Indeed, suppose that a strategy of the form $\hat{\sigma}$ is optimal. The payoff of $\hat{\sigma}$ is given by

$$v_0 + \delta \mathbb{E}_{\hat{\sigma}} \left[\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t-1} \max\{0, r + \hat{\alpha}_{\tau-t-1}(q - r) - (1 - p)\} + \delta^{\hat{t}-t} \frac{1 - \delta^{N-\hat{t}}}{1 - \delta} v_0 \right],$$

where $\hat{\alpha}_n$ denotes the agent's (random) reputation following n observations after period $N - k$, and the expectation is taken with respect to $\hat{t}$ and the $\hat{\alpha}_n$. Now, we consider the strategy σ, which modifies $\hat{\sigma}$ so that the action prescribed by $\hat{\sigma}$ is taken one stage earlier (and the agent is not employed for sure in the last period N). The payoff of this strategy σ is given by

$$\mathbb{E}_{\hat{\sigma}} \left[\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t-1} \max\{0, r + \hat{\alpha}_{\tau-t-1}(q - r) - (1 - p)\} + \delta^{\hat{t}-t} \frac{1 - \delta^{N-\hat{t}+1}}{1 - \delta} v_0 \right].$$

Subtracting the payoff of $\hat{\sigma}$ from the payoff of σ gives us

$$\mathbb{E}_{\hat{\sigma}} \left[\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t-1} \left(\max\{0, r + \hat{\alpha}_{\tau-t-1}(q - r) - (1 - p)\} - v_0 \right) \right] (1 - \delta).$$

[12] The proof of Claim 1 is an adaptation of the proof of Theorem 5.2.2 in Berry and Fristedt (1985).

As $\hat{\sigma}$ is optimal, its payoff is at least as large as that from never acquiring the signal, which is $\frac{1-\delta^{N-t+1}}{1-\delta}v_0$. As $\frac{1-\delta^{N-t+1}}{1-\delta} = \sum_{t=t+1}^{\hat{t}} \delta^{\tau-t-1} + \delta^{\hat{t}-t}\frac{1-\delta^{N-\hat{t}+1}}{1-\delta}$ for all $\hat{t} \in \{t+1, t+2, \cdots, N\}$, we can write

$$v_0 + \delta\mathbb{E}_{\hat{\sigma}}\left[\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t-1}\max\{0, r+\hat{\alpha}_{\tau-t-1}(q-r)-(1-p)\} + \delta^{\hat{t}-t}\frac{1-\delta^{N-\hat{t}}}{1-\delta}v_0\right]$$
$$\geq v_0 + \delta\frac{1-\delta^{N-t+1}}{1-\delta}v_0 = \mathbb{E}_{\hat{\sigma}}\left[\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t-1}v_0 + \delta^{\hat{t}-t}\frac{1-\delta^{N-\hat{t}}}{1-\delta}v_0\right],$$

where the inequality follows from the optimality of $\hat{\sigma}$. Rearranging, we get $\mathbb{E}_{\hat{\sigma}}\left[\sum_{\tau=t+1}^{\hat{t}} \delta^{\tau-t-1}\left(\max\{0, r+\hat{\alpha}_{\tau-t-1}(q-r)-(1-p)\} - v_0\right)\right] \geq 0$. Thus, we have shown that the payoff of σ is at least as large as the payoff of $\hat{\sigma}$, and the induction step is complete.

We have thus proved Claim 1 for any arbitrary finite horizon $N \in \mathbb{N}$. Now, let us assume that $N = \infty$. Let σ be the optimal strategy that fires the agent if and only if it is *strictly* optimal to do so. On account of discounting, the game is continuous at infinity, and thus, whenever σ prescribes that the agent be fired at a given stage t, firing him is uniquely optimal at t for finite-horizon approximations with a long enough horizon. For each such approximation, not hiring him in all periods following t is optimal. By continuity of payoffs at infinity, this implies that not hiring him after time t is optimal for $N = \infty$ as well.

Claim 2: The payoff of hiring the agent in any period $t \in \{1, \cdots, N\}$ is non-decreasing in the belief α. Thus, there exists an optimal policy with a threshold structure in which the agent is never rehired once fired.

First, let $N < \infty$. We proceed by induction over k. In the last period, i.e. for $k = 0$, the agent is hired if and only if the benefit from following his signal exceeds the opportunity cost of hiring him, v_0. Thus, $V_N(\alpha) = \max\{v_0, r+\alpha(q-r)-(1-p)\}$. This is clearly non-decreasing in α. Furthermore, the first-best policy is to employ the agent if and only if $\alpha \geq \frac{v_0+1-p-r}{q-r} =: \bar{\alpha}(0)$.

Now, suppose that V_τ is non-decreasing for all $\tau \in \{t+1, t+2, \cdots, N\}$. We now show that V_t is non-decreasing as well. Using Claim 1, we have that

$$V_t(\alpha) = \max\left\{v_0\frac{1-\delta^{N-t+1}}{1-\delta}, \max\{0, r+\alpha(q-r)-(1-p)\} + \delta\mathbb{E}[V_{t+1}(\alpha_{t+1})|\alpha_t = \alpha]\right\}.$$

We have that $\max\{0, r+\alpha(q-r)-(1-p)\}$ is non-decreasing. It remains to be shown that $\delta\mathbb{E}[V_{t+1}(\alpha_{t+1})|\alpha_t = \alpha]$ is non-decreasing in α. This follows from the fact that V_{t+1} is non-decreasing (by the induction hypothesis) and that the distribution of α_{t+1} conditional on $\alpha_t = \alpha'$ first-order stochastically dominates the distribution of α_{t+1} conditional on $\alpha_t = \alpha''$, for all $\alpha' > \alpha''$. It thus follows that V_t is non-decreasing as well.

Now, suppose that, in period t, hiring the agent is optimal for a belief $\bar{\alpha}$. It follows that

$$\max\{0, r+\bar{\alpha}(q-r)-(1-p)\} + \delta\mathbb{E}[V_{t+1}(\alpha_{t+1})|\alpha_t = \bar{\alpha}] \geq v_0\frac{1-\delta^{N-t+1}}{1-\delta}.$$

Since the left-hand side is non-decreasing in $\bar{\alpha}$, it follows that hiring the agent is optimal for all beliefs $\alpha > \bar{\alpha}$.

Thus, for finite horizons, we have shown the existence of an optimal policy with a threshold structure in which the agent is never rehired once fired; i.e. there exists a sequence of thresholds $(\bar{\alpha}(N-t))_{t=1}^{N}$ such that the agent is employed in period t if and only if $\alpha_t \geq \bar{\alpha}(N-t)$.

Now, let $N = \infty$. Let $\hat{\sigma}$ be an arbitrary strategy that has the structure of Claim 1. If $\hat{\sigma}$ does not call for hiring the agent in the current period, the agent will never be hired in the future. The payoff from $\hat{\sigma}$ is thus non-decreasing in α in this case.

Suppose then that $\hat{\sigma}$ calls for hiring the agent in the current period. In each period, the strategy $\hat{\sigma}$ maps the number of previous successes and failures into a decision of either hiring or not hiring the agent, and thus leads to a distribution over stopping times $\hat{t}+1$. For any given fixed $\hat{t} \in \{0, 1, \cdots\} \cup \{\infty\}$, the payoff given an initial belief of α can be written as

$$V_{\hat{t}}(\alpha) = \sum_{t=0}^{\hat{t}} \delta^t \left(\mathbb{E}\left[\max\{0, r + \hat{\alpha}_t(q-r) - (1-p)\} | \hat{\alpha}_0 = \alpha\right] - v_0 \right) + \frac{v_0}{1-\delta},$$

where the expectation is taken with respect to $\{\hat{\alpha}_t\}_{t=0}^{\hat{t}}$. For any $t \in \{0, 1, \cdots, \hat{t}\}$, $\max\{0, r + \hat{\alpha}_t(q-r) - (1-p)\}$ is non-decreasing in $\hat{\alpha}_t$. As, for any t, the distribution of $\hat{\alpha}_t$ conditional on $\hat{\alpha}_0 = \alpha'$ first-order stochastically dominates the distribution of $\hat{\alpha}_t$ conditional on $\hat{\alpha}_0 = \alpha''$, for all $\alpha' > \alpha''$, this implies that $V_{\hat{t}}(\alpha)$ is non-decreasing, for any given $\hat{t} \in \{0, 1, \cdots\} \cup \{\infty\}$. As $V(\hat{\sigma}; \alpha)$, the decision maker's payoff from the strategy $\hat{\sigma}$, satisfies $V(\hat{\sigma}; \alpha) = \mathbb{E}_{\hat{\sigma}}\left[V_{\hat{t}}(\alpha)\right]$, where the expectation is taken with respect to the stopping time $\hat{t}+1$, it follows that $V(\hat{\sigma}; \cdot)$ is non-decreasing. Hence the decision maker's payoff $V(\alpha) := \sup_{\hat{\sigma}} V(\hat{\sigma}; \alpha)$ is non-decreasing. By the Principle of Optimality, we can furthermore write

$$V(\alpha) = \max\left\{ \frac{v_0}{1-\delta}, \max\{0, r + \alpha(q-r) - (1-p)\} + \delta \mathbb{E}[V(\alpha_{t+1}) | \alpha_t = \alpha] \right\}.$$

Now, suppose that it is optimal to hire the agent at a given belief $\bar{\alpha}$. Then,

$$\max\{0, r + \bar{\alpha}(q-r) - (1-p)\} + \delta \mathbb{E}[V(\alpha_{t+1}) | \alpha_t = \bar{\alpha}] \geq \frac{v_0}{1-\delta}.$$

Since $V(\alpha)$ is non-decreasing, and the distribution of α_{t+1} conditional on $\alpha_t = \alpha'$ first-order stochastically dominates the distribution of α_{t+1} conditional on $\alpha_t = \alpha''$ for all $\alpha' > \alpha''$, the left-hand side of this inequality is non-decreasing in $\bar{\alpha}$. It thus follows that it is optimal to hire the agent for all beliefs $\alpha > \bar{\alpha}$ as well. Thus, we have shown the existence of an optimal policy with a threshold structure in which the agent is never rehired once fired for an infinite horizon as well.

Claim 3: The thresholds $\bar{\alpha}(N-t)$ are non-decreasing in t.

Consider an optimal policy which satisfies Claims 1 and 2, and which is characterized by the sequence of thresholds $(\bar{\alpha}(N-t))_{t=0}^{N}$. Suppose to the contrary that there exists a period t such that $\bar{\alpha}(N-t) > \bar{\alpha}(N-t-1)$ for our optimal policy. Consider a belief $\alpha_t \in (\bar{\alpha}(N-t-1), \bar{\alpha}(N-t))$. Then, in our optimal policy, the expert is not hired in period t, but is again hired in period $t+1$. This contradicts Claim 1.

Claim 4: For $N = \infty$, there exists an optimal policy characterized by a fixed threshold $\bar{\alpha}$.

This immediately follows from Claims 1-2 and the fact that the decision maker's problem for a given belief α is the same in any two periods t_1, t_2. Therefore, if a threshold of $\bar{\alpha}$ is optimal at a given time t_1, it is optimal at all other times t_2 as well. ∎

Proof of Theorem 2

If the first-best outcome is never to employ the agent, this outcome is trivially incentive compatible.

We now consider the more interesting case in which in any first-best outcome the agent is employed in the first period. If $v_0 = 0$, there is a first-best outcome in which the agent is employed in each period. This outcome, again, is trivially incentive compatible.

Let now $v_0 > 0$ and consider the case with $q = 1$. The first-best outcome employs the agent in the first period and continues to employ the agent in the subsequent periods if and only if all his previous signals are correct.[13] We construct an equilibrium that implements this first-best outcome as follows. When the agent has not previously deviated, he accepts employment whenever doing so is consistent with the first-best outcome; whenever he is employed and has not previously deviated, he reports his information truthfully. If the agent deviates, he follows an optimal continuation strategy given the principals' behavior. As any history of reports and state realizations can occur on the path of play, each principal's belief is pinned down by Bayes' rule after any history.

By assumption, the period t-principal's wage offer is $v_t = v(\alpha_t)$, i.e., $v_t = v_t^*$ if, in all previous periods, the agent's reports have been correct, and $v_t = 0$ otherwise.

Furthermore, the agent's employment decision on the equilibrium path is optimal since the agent is the residual claimant of the surplus and it is optimal to employ the agent in the first-best outcome.

We now show that reporting truthfully is optimal on the equilibrium path. Whenever the agent deviates from his reporting strategy on the equilibrium path in period $t-1$, he either makes an incorrect report, in which case his continuation payoff is v_0 per period, or he privately learns that he is incompetent because his untruthful report was correct. In the latter case, the agent's expected continuation payoff in period t, divided by δ, is bounded from above by

$$\underline{v}_t^0 := \sum_{\tau=t}^{N} \delta^{\tau-t} \left((1-p)^{\tau-t} p + (1-(1-p)^{\tau-t}) v_0 \right).$$

To understand this expression, recall that p is the maximal feasible period wage and observe that given the strategies of the principals an incompetent agent reveals himself, and obtains the continuation payoff of v_0 in each period, with a probability of at least p. Indeed, the agent will continue to receive a positive wage if and only if his report is correct. Regardless of the reporting strategy, an incompetent agent's report can be correct with a probability of at most $1-p$.

On the equilibrium path, the agent's expected continuation payoff in period t, divided by δ and evaluated at the beginning of period $t-1$, is bounded from below by

$$\overline{v}_t^* := \max_{t' \geq t} \alpha_{t-1}^* \sum_{\tau=t'}^{N} \delta^{\tau-t} v_{t'}^* = \max_{t' \geq t} \delta^{t'-t} \frac{1-\delta^{N+1-t'}}{1-\delta} \alpha_{t-1}^* v_{t'}^*.$$

Given these bounds, the agent's strategy is a best reply on the equilibrium path in period t if

$$\overline{v}_{t+1}^* \geq \underline{v}_{t+1}^0. \tag{A.1}$$

Let $\hat{t}$ be the minimal value of t such that

$$\alpha_t^* v_{t+1}^* > (1-p)p + p v_0.$$

Note that the left-hand side of the inequality is (weakly) increasing and, as $\lim_{t\to\infty} \alpha_t^* = 1$ and $\lim_{t\to\infty} v_{t+1}^* = p$, converging to p as $t \to \infty$. Since $v_0 < \alpha p$, $\hat{t}$ is finite.

It follows then that (A.1) is satisfied for all $t \geq \hat{t}$. Let now $t < \hat{t}$. Note that

$$\lim_{N\to\infty} \underline{v}_t^0 \quad = \quad \underline{v} := \frac{v_0}{1-\delta} + \frac{p - v_0}{1-\delta(1-p)},$$

[13]Recall that $q = 1$ and, therefore, an incorrect signal reveals the agent's incompetence. Note that the first-best rule continues to have this structure if N increases.

whereas

$$\lim_{N\to\infty} \overline{v}_t^* \quad = \quad \lim_{N\to\infty} \max_{t'\geq t} \delta^{t'-t} \frac{1-\delta^{N+1-t'}}{1-\delta} \alpha_{t-1}^* v_{t'}^*$$

Furthermore, as $\delta \to 1$, both expressions diverge to infinity and their ratio converges to

$$\gamma := \frac{v_0}{\alpha_{t-1}^* p} < 1, \tag{A.2}$$

since $\alpha p > v_0$ and $v_t^* \to p$ as $N \to \infty$. Hence,

$$\lim_{\delta\to 1} \lim_{N\to\infty} \left(\overline{v}_{t+1}^* - \underline{v}_{t+1}^0\right) = \infty.$$

Defining $\check{\delta}_0$ to be the infimum value of δ among all δ for which

$$\lim_{N\to\infty} \overline{v}_{t+1}^* > \lim_{N\to\infty} \underline{v}_{t+1}^0. \tag{A.3}$$

for all $t+1 < \hat{t}$ completes the proof for $q = 1$. All these arguments trivially extend to the case of $N = \infty$.

From (A.3), we know that all incentive constraints hold with slackness for $q = 1$ for any given $N > N_\delta$ and $\delta > \check{\delta}_0$. Now, for any fixed finite $N > N_\delta$ and $\delta > \check{\delta}_0$, the parties' payoffs in the first-best outcome are Lipschitz continuous in q in a neighborhood of $q = 1$. Furthermore, by assumption, it is strictly optimal to employ the agent in the first-best outcome if $q = 1$. Therefore, there exists a $q_{N,\delta} < 1$ such that, for $q \geq q_{N,\delta}$, the first-best outcome employs the agent in the first period, the principals' wage offers are 0 after the first report that does not match the state, and the agent finds it optimal to report his information truthfully. ∎

Proof of Proposition 3

Choose $\bar{\delta}_0 \in (0,1)$ so that, for all discount factors $\delta \in (\bar{\delta}_0, 1)$, the first best be incentive compatible in the infinite-horizon game. (The proof of Theorem 2 shows that such a $\bar{\delta}_0$ exists.)

Consider the game with horizon $\hat{N}$ and discount factor $\hat{\delta} \in (\bar{\delta}_0, 1)$, such that (A.1) holds in all periods $t \in \{1, \cdots, \hat{N}\}$. It is sufficient to show that (A.1) continues to hold for all discount factors $\delta \in (\hat{\delta}, 1)$, for the same horizon $\hat{N}$. If $\hat{N} = \infty$, there is nothing left to show. We therefore assume that $\hat{N} < \infty$. We fix an arbitrary period t. As (A.1) holds in all periods, we have that $\sum_{\tau=k}^{\hat{N}} \hat{\delta}^{\tau-k} g(\tau) \geq 0$ for all $k \in \{t, \cdots, N\}$, where g is defined by

$$g(\tau) := \begin{cases} -v_0 - (p-v_0)(1-p)^{\tau-t} & \text{if} \quad \tau < t' \\ \alpha_{t-1}^* v_{t'}^* - v_0 - (p-v_0)(1-p)^{\tau-t} & \text{if} \quad \tau \geq t' \end{cases},$$

where $t' \in \arg\max_{\tilde{t}\geq t} \alpha_{t-1}^* \sum_{\tau=\tilde{t}}^{\hat{N}} \hat{\delta}^{\tau-t} v_{\tilde{t}}^*$.

Now, we consider an arbitrary $\delta > \hat{\delta}$, and shall show that $\sum_{\tau=k}^{\hat{N}} \hat{\delta}^{\tau-k} g(\tau) \geq 0$ for all $k \in \{t, \cdots, N\}$ implies $\sum_{\tau=k}^{\hat{N}} \delta^{\tau-k} g(\tau) \geq \sum_{\tau=k}^{\hat{N}} \hat{\delta}^{\tau-k} g(\tau) \geq 0$ for all $k \in \{t, \cdots, N\}$. This is sufficient to show that (A.1) will continue to hold in all periods $k \in \{t, \cdots, \hat{N}\}$ for $\delta > \hat{\delta}$ as well.

The proof is by induction over k. For $k = \hat{N}$, the claim is equivalent to $g(\hat{N}) \geq 0$, which is true as (A.1) holds in period $t = N$. We posit as our induction hypothesis that $\sum_{\tau=k}^{\hat{N}} \delta^{\tau-k} g(\tau) \geq$

$\sum_{\tau=k}^{\hat{N}} \hat{\delta}^{\tau-k} g(\tau) \geq 0$ for all $k \in \{\bar{k}, \cdots, \hat{N}\}$. Now, $\sum_{\tau=\bar{k}-1}^{\hat{N}} \delta^{\tau-\bar{k}+1} g(\tau) = g(\bar{k}-1) + \delta \sum_{\tau=\bar{k}}^{\hat{N}} \delta^{\tau-\bar{k}} g(\tau)$. It remains to show that

$$g(\bar{k}-1) + \delta \sum_{\tau=\bar{k}}^{\hat{N}} \delta^{\tau-\bar{k}} g(\tau) \geq g(\bar{k}-1) + \hat{\delta} \sum_{\tau=\bar{k}}^{\hat{N}} \hat{\delta}^{\tau-\bar{k}} g(\tau) \geq 0$$

$$\iff \sum_{\tau=\bar{k}}^{\hat{N}} \delta^{\tau-\bar{k}} g(\tau) \geq \frac{\hat{\delta}}{\delta} \sum_{\tau=\bar{k}}^{\hat{N}} \hat{\delta}^{\tau-\bar{k}} g(\tau) \geq 0,$$

which is verified, as, by our induction hypothesis, $\sum_{\tau=\bar{k}}^{\hat{N}} \delta^{\tau-\bar{k}} g(\tau) \geq \sum_{\tau=\bar{k}}^{\hat{N}} \hat{\delta}^{\tau-\bar{k}} g(\tau) \geq 0$, and $\delta > \hat{\delta}$. ∎

B Counter-Example

Let $\alpha = 5/12$, $p = 3/7$, $q = 9/10$, $r = 1/2$, and $\delta = 1$.

1. Let $N = 2$. If v_0 is positive and sufficiently small, the first-best outcome employs the agent in period 2 if and only if his signal is correct in period 1.

2. Let $N = 3$. If v_0 is positive and sufficiently small, the first-best outcome always employs the agent in period 2, and employs him in period 3 if and only if his signal was correct at least once in the previous two periods.

   This rule is *not* incentive compatible. Indeed, for the parties to benefit from the agent's expertise in the last period, the agent's wage offer in that period should be larger than v_0. Yet, if this were true, the agent's best response after an incorrect signal in period 1 is to disregard his signal and report 0 in period 2. In equilibrium, the public belief about the agent's competence cannot be updated based on the second-period reports. This depresses the last period wage offer and makes it impossible to condition employment on the signal realization in the second period. As a consequence, either the agent will not be employed in the last period, because of a lower wage offer, when the first-best outcome mandates employment or the agent will be employed in the last period when he should not be because the signal realization in the second period is concealed.

References

Akerlof, G. A. (1970): "The Market For "Lemons": Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), 488–500.

Berry, D. A. and B. Fristedt (1985): *Bandit Problems: Sequential Allocation of Experiments.* London: Chapman & Hall.

Board, S. and M. Meyer-ter-Vehn (2013): "Reputation for Quality," *Econometrica*, 81(6), 2381–2462.

Bonatti, A. and J. Hörner (2017): "Career Concerns with Exponential Learning," *Theoretical Economics*, 12, 425–475.

Bourjade, S. and B. Jullien (2011): "The Roles of Reputation and Transparency on the Behavior of Biased Experts," *RAND Journal of Economics*, 42, 575–594.

Dasgupta, A. and A. Prat (2006): "Financial Equilibrium with Career Concerns," *Theoretical Economics*, 1, 67–93.

Dasgupta, A. and A. Prat (2008): "Information Aggregation in Financial Markets with Career Concerns," *Journal of Economic Theory*, 143(1), 83–113.

Effinger, M. R. and M. Polborn (2001): "Herding and Anti-Herding: A Model of Reputational Differentiation," *European Economic Review*, 45(3), 385–403.

Ely, J. C. and J. Välimäki (2003): "Bad Reputation," *Quarterly Journal of Economics*, 118(3), 785–814.

Fama, E. F. (1980): "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88(2), 288–307.

Holmström, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169–182.

Jullien, B. and I. Park (2014): "New, Like New, or Very Good? Reputation and Credibility," *Review of Economic Studies*, 81, 1543–1574.

Leaver, C. (2009): "Bureaucratic Minimal Squawk Behavior: Theory and Evidence from Regulatory Agencies," *American Economic Review*, 99(3), 572–607.

Levy, G. (2004): "Anti-Herding and Strategic Consultation," *European Economic Review*, 48(3), 503–525.

Levy, G. (2007): "Decision Making in Committees: Transparency, Reputation and Voting Rules," *American Economic Review*, 97(1), 150–168.

Morgan, J. and P. C. Stocken (2003): "An Analysis of Stock Recommendations," *RAND Journal of Economics*, 34(1), 183–203.

Morris, S. (2001): "Political Correctness," *Journal of Political Economy*, 109(2), 231–265.

Ottaviani, M. and P. N. Sørensen (2006a): "Professional Advice," *Journal of Economic Theory*, 126(1), 120–142.

Ottaviani, M. and P. N. Sørensen (2006b): "Reputational Cheap Talk," *RAND Journal of Economics*, 37(1), 155–175.

Pavesi, F. and M. Scotti (2014): "Experts, Conflicts of Interest, and Reputation for Ability," *International Journal of Economic Theory*, 10(2), 219-233.

Prat, A. (2005): "The Wrong Kind of Transparency," *American Economic Review*, 95(3), 862–877.

Prendergast, C. and L. Stole (1996): "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 104(6), 1105–1134.

Scharfstein, D. S. and J. C. Stein (1990): "Herd Behavior and Investment," *American Economic Review*, 80(3), 465–479.

Trueman, B. (1994): "Analyst Forecasts and Herding Behavior," *Review of Financial Studies*, 7(1), 97–124.