Should the Flatterers be Avoided?*

Nicolas Klein† Tymofiy Mylovanov‡

This version: February 15, 2011

Preliminary and Incomplete.
In particular, c.f. footnote 7.

Abstract

We analyze a dynamic career concerns game between an expert and a decision maker. In each period, the decision maker has the option of obtaining cheap-talk advice from the expert, who is merely interested in his continued employment. The expert's quality is initially unknown to both parties. The incentive problem is that the expert might attempt to avoid appearing uninformed by concealing his true opinion if it contradicts the predominant view. Nevertheless, if a competent expert never makes mistakes, the fully revealing first-best outcome can be implemented provided the time horizon is sufficiently long. For shorter time horizons, the optimal equilibrium in this case features a grace period during which the expert accumulates some private information about his quality.

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

JEL Classification Numbers: C73, D83.

---

*We are grateful to Dirk Bergemann, Matthias Fahn, Johannes Hörner, Lucas Maestri, George Mailath, Benny Moldovanu, Stephen Morris, Sven Rady, Phil Reny, and Satoru Takahashi for useful comments and discussions. The first author is particularly grateful to the Department of Economics at Yale University and the Cowles Foundation for Research in Economics for an extended stay, during which this paper took shape. Financial support from the National Research Fund, Luxembourg, and the German Research Fund through SFB TR 15 is gratefully acknowledged.

†University of Bonn, Lennéstr. 37, D-53113 Bonn, Germany; email: kleinnic@yahoo.com.

‡Department of Economics, Penn State University; email: txm41@psu.edu.

> THE choice of servants is of no little importance to a prince, and they are good or not according to the discrimination of the prince. And the first opinion which one forms of a prince, and of his understanding, is by observing the men he has around him; and when they are capable and faithful he may always be considered wise, because he has known how to recognize the capable and to keep them faithful.
>
> Nicolo Machiavelli, "The Prince,"
> Ch. XXII "Concerning The Secretaries of Princes"

> Because there is no other way of guarding oneself from flatterers except letting men understand that to tell you the truth does not offend you; but when every one may tell you the truth, respect for you abates.
>
> Nicolo Machiavelli, "The Prince,"
> Ch. XXIII "How Flatterers Should Be Avoided"

1 Introduction

To most people, Machiavelli's wariness of flatterers will come as no surprise. Flatterers merely reenforce existing preconceptions; they will not tell the prince what they know through candid unvarnished advice, and hence their knowledge will be lost to the prince. However, we shall make the point that there is a cost to avoiding flatterers also; as a result, flattery by advisors should sometimes be encouraged.

We study a model of a decision maker (she) consulting an expert (he) whose quality is uncertain and is unknown to either of them. The expert is strategic and is only interested in his continued employment. The core issue in the model is the following: If the expert were not strategic, the decision maker would like to employ the expert if and only if the expert continued to prove his competence. Yet, if this decision rule were in fact followed, the expert might have incentives to suppress a priori unlikely information to maximize his chances of appearing competent; this would slow down learning about the expert's quality and thus render his advice less valuable. In this paper, we analyze how best to benefit from the advice of the capable and how to avoid being corrupted by the flattery of the incompetent. Counterintuitively, the optimal equilibria can feature grace periods during which the expert suppresses his information and provides useless advice.

The pitfalls of inadequate advice are legion. Think of King David's son Absalom, who, spurning Ahithophel's wise counsel, met his untimely demise in battle, whilst, offended by the

slight, Ahithophel committed suicide (II Sam 15-17). Another well-known ill-fated advisor was Thomas Cromwell, Earl of Essex, who advised Henry VIII to marry Anne of Cleves. The eventual failure of this marriage, the king's fourth out of a total of six, contributed to Cromwell's execution.

For a more mundane and more specific motivation consider, for instance, a legislator who can consult a pollster to find out whether the people she represents would prefer a vote in favor of, or against, a particular bill under consideration. Not fully certain about the quality of his research, the pollster might decide to play it safe and distort his report toward the commonly known ideological bent of the legislator's district. The legislator's goal meanwhile is to make the best possible use of the advice she gets.

More broadly, our model applies to many instances when a decider who faces a sequence of multiple decision problems can elect to seek the help of outside advisors whose quality is initially unknown: A firm trying to gauge the demand for a new product may look toward outside consultants for advice. People filing their income tax returns may hire a tax consultant. Or a college student having set aside an afternoon to study for a test may decide to spend the time with a tutor.

We study the simplest model that formalizes the expert's reputational concerns in an explicitly dynamic setting. In our model, there are $N \geq 2$ periods. In each period, the decision maker chooses a policy. The optimal policy is uncertain and is described by the state of nature which is a binary random variable which is iid across periods. The decision maker's payoff is publicly revealed at the end of the period.

If consulted, the expert observes a binary noisy signal about the realization of the state, upon which he sends a cheap-talk message to the decision maker. The quality of the signal is initially unknown to both parties. The low-quality signal is uninformative and is always equally likely to be correct or incorrect. The high-quality signal is informative and is correct with a time-invariant probability. The expert's objective is to be consulted as often as possible.

In order to focus on the expert's incentives and to clarify the core intuition behind our main insights, we restrict the decision maker's behavior and require that she terminate the expert if there is no value in continuing to employ him. This restriction could be viewed as a reduced-form representation of behavior in a richer model in which the decision maker has limited commitment power and incurs an opportunity cost of employing the expert.

Our model belongs to the class of 'bad reputation' models: The expert's concern about appearing to be competent creates incentives for him to distort his reports. It is related to Ottaviani and Sørensen (2006a, 2006b) who explore the effect of this type of reputation

concerns on the reporting strategy of the expert in *a single-decision environment*. The important message of their work is that, under weak conditions, truthful reporting is not feasible and the information that can be communicated in equilibrium, if any, is limited.[1]

In *multi-decision environments*, such as our model, there is an additional *good news* effect. The expert's private information about whether his signal was correct impacts his evaluation of the expected continuation payoff after different public histories through his belief about his competence. By contrast with single-decision environments, the good news effect can make truthful reporting feasible.

Indeed, consider a hypothetical first-best environment in which the decision maker can observe the expert's signals. In this environment, the first-best decision rule would prescribe continuing to employ the expert until a certain number $\kappa \geq 0$ of incorrect signals was observed.[2] We show that if the competent expert never makes mistakes, the good news effect restores incentives for the expert to report truthfully in the first-best decision rule as the number of periods grows sufficiently large (Proposition 4.2). This result is based on the observation that after a correct report the difference between the continuation payoff after the expert has told the truth and the continuation payoff after he has lied is increasing and diverging in N, while the expert's continuation payoff after a mistake is always zero.

If the competent expert occasionally makes mistakes, however, increasing the time horizon does not suffice to resolve the incentive problem. In fact, the opposite can happen. As we show in an example an increase in the number of periods can make the first-best outcome unattainable. Proposition 4.3 establishes that the first-best decision rule is not incentive compatible if the number of periods is sufficiently large.

The introductory quote by Machiavelli suggests that the decision maker should be concerned with creating incentives for the expert to avoid flattery and to tell the truth whenever his private information disagreed with the predominant view. Indeed, the issue of truthful information transmission has been one of the main questions in the literature on cheap-talk communication. We show that even for shorter time horizons for which the first-best outcome cannot be attained, truthful reporting of the expert's information can be supported in equilibrium (Proposition 5.1). This equilibrium conditions the expert's retainment both on the public belief about his competence as well as on some payoff-irrelevant features of

[1] Similar results are obtained in Morris (2001) and Morgan & Stocken (2003) who study cheap talk models in which the (bad) reputation concerns pertain to the preferences of the expert rather than his competence.

[2] There is a value in continuing to learn about the expert's quality if and only if there is a chance that the expert's advice will be relevant for the policy choice in the last period. A necessary and sufficient condition for the optimal choice of policy in the last period to match the expert's signal is that the expert has observed κ or fewer incorrect reports in total.

the reporting history. In equilibrium, the incentives for the expert to report his information are created by a threat of a higher probability of termination after the *a priori* more likely states.

Intuitively appealing though it may be, truthful reporting ought not to be an end in itself for the decision maker. If the competent expert never makes mistakes, the optimal equilibrium features an initial grace regime during which the expert babbles and accumulates some private information about his quality, which is followed by a monitoring regime in which the expert is fired whenever his report disagrees with the optimal policy (Proposition 5.3). At the end of the grace regime, an expert who has not observed any incorrect signals is sufficiently optimistic about his competence to report his signal in the monitoring regime. On the other hand, an expert who has learned his incompetence during the grace regime will disregard his signal and send but uninformative messages, which coincide with the policy that is first-best optimal conditional on the expert's being of poor quality.

Thus, in the optimal equilibrium, the decision maker's policy choices coincide with those she would make in the first-best environment, although she might end up employing an incompetent expert for awhile. By contrast, in the fully revealing equilibrium, the implementation of truthtelling in the early periods requires the firing of a competent expert with non-negligible probability; this is costly as the decision maker will now with some probability forgo the benefit of competent advice in the future. One particular implication of these considerations is that, surprisingly, the agency problem in our model increases the expected tenure of the expert.

In reference to the question in the title, our results suggest that unless the first-best decision rule is implementable, avoiding flatterers altogether is bad judgement.

The rest of this paper is structured as follows: Section 2 reviews some related literature; Section 3 presents the model, and introduces necessary notation; in Section 4, we analyze the first-best decision rules, whereas the second-best decision rules are studied in Section 5; Section 6 concludes. The proofs omitted in the main text are provided in the appendix.

2 Related Literature

As discussed in the introduction, our model is most related to the literature on reputational cheap talk (Ottaviani and Sørensen (2006a, 2006b), Morris (2001), and Morgan & Stocken (2003)). In this section, we discuss other related literature.

Our decision maker's problem is akin to that of an economic agent operating a two-armed

bandit machine with one safe arm, whose expected payoff is known, and one risky arm, whose expected payoff is initially unknown but can potentially be learnt through use over time. Bandit models have been used in economics, at least since Rothschild's (1974) discrete-time model, to analyze the trade-off between experimentation and exploitation, providing insights into agents' incentives to forgo current payoffs in exchange for information which can potentially be parlayed into better decisions, and thus higher payoffs, come tomorrow.[3] The same trade-off is central to our paper, with the additional twist that our "risky arm" is acting strategically, in the sense that the expert will choose his messages with a view to maximizing his expected duration of employment.

The single-agent two-armed bandit decision problem has been analyzed as early as the 1950s (cf. e.g. Bradt, Johnson, Karlin, 1956). The analysis of the strategic interplay between several agents operating replica bandits (cf. Bolton & Harris, 1999, 2000, Keller, Rady, Cripps, 2005, Keller & Rady, 2010), negatively correlated bandits (cf. Klein & Rady, 2010), or three-armed bandits (cf. Klein, 2010a), is much more recent, by contrast.

"Strategic risky arms", in the sense that the party about whom information is gathered is also taking an action, have been studied by Bergemann & Välimäki (1996, 2000) and Bar-Isaac (2003). Bergemann & Välimäki (1996) investigate the case of a *single* buyer desiring to purchase a good from any one of multiple firms, which produce at different, and initially unknown, qualities. They show that, with firms pricing the good strategically, experimentation is at efficient levels in any Markov perfect equilibrium. In Bergemann & Välimäki (2000), by contrast, they show that with multiple buyers and two firms, the firms are subsidizing experimentation to the point that it reaches inefficiently high levels in equilibrium. In Bar-Isaac (2003), a seller of a good of initially unknown quality faces a sequence of buyer pairs, who, in each period the seller (strategically) decides to sell, Bertrand-compete for the good. After the transaction, the quality of the good is publicly observed, and beliefs about the seller's quality are updated. In contrast to these papers, we investigate

Our model is also related to the literature on dynamic agency under gradual learning, which has e.g. been investigated by Bergemann & Hege (1998, 2005) and Hörner & Samuelson (2009), who examine a venture capitalist's optimal provision of funds to an entrepreneur, who might divert the funds toward his private ends. As is the case in our model, the accumulation of private information by an agent subsequent to a deviation from the equilibrium path provides an interesting dynamic aspect to the strategic interaction. However, our model is probably closest to Gerardi & Maestri (2008) who analyze the case of an expert who, in order to receive an informative signal about the decision-relevant state of nature, has to incur some private effort costs. After choosing whether or not to expend effort on acquiring the

[3] cf. Bergemann & Välimäki (2008) for an overview of this literature.

signal, the expert then sends the decision maker a cheap-talk message about the state of the world. Our model crucially differs from all these papers, though, in that we are investigating a problem of *adverse selection*, or *hidden information*, rather than *moral hazard*: Our expert privately observes a signal without having to incur any effort costs; it is rather his innate quality, which determines the precision of his signal, which is uncertain at first.[4] Over the course of the interaction, the parties gradually update their beliefs about the expert's type; as he wants to continue to be thought of highly in order to be re-employed, our expert will have incentives strategically to manipulate the decision maker's learning process.

3 Model Setup

There is a decision maker and an expert who live for N periods, $t = 1, \ldots, N$.[5] In each period, the decision maker has to choose a policy from the set $\{0, 1\}$. The optimal policy is uncertain and is described by the state $\omega_t \in \{0, 1\}$. The value of ω_t is distributed independently across periods and is equal to 1 with a probability of $p \in (0, 1/2)$, which is common knowledge. The decision maker obtains a period payoff of 1 whenever the policy matches the state and 0 otherwise; there is no discounting.

In each period, the decision maker can either consult an expert or rely on her prior in choosing the policy. If consulted, the expert observes a non-verifiable signal $\tilde{s} \in \{0, 1\}$ about the realized state. The distribution of $\tilde{s}$ is iid across periods and depends on the quality of the expert, which is initially unknown to both players. With probability $\alpha \in (0, 1)$ the expert is competent and his signal reveals the optimal policy with probability $q \in (1 - p, 1]$. With complementary probability, the expert is incompetent, in which case $\tilde{s}$ is distributed uniformly and independently of the state.

We denote by α_t the decision maker's belief about the expert's competence at the beginning of period t; we refer to it as the expert's *reputation*. The expert's belief is denoted by $\hat{\alpha}_t$. This belief could well differ from the decision maker's because the expert has the benefit of privately knowing the signals he has observed. We use β_t and $\hat{\beta}_t$ to denote respectively the decision maker's and the expert's belief that the signal in period t is correct.

We introduce *career concerns* on the expert's part into our dynamic setting by assuming that it is the expert's objective to be employed for as many periods as possible. Hence, we

[4] While initially there is incomplete but symmetric information as to the expert's quality, private information may arise endogenously both on and off the equilibrium path, as we shall see.

[5] Alteratively, we could assume infinitely many periods with discounting. Yet, end-of-game effects are not much of a concern in our model, so that we would expect similar qualitative results with infinitely many periods.

assume that the expert gets a payoff of 1 per period when he is employed and 0 otherwise. Again, there is no discounting. To rule out uninteresting cases, we impose the following

Assumption 3.1 *It is commonly believed that* $\beta_1 := \alpha q + (1-\alpha)/2 < 1 - p$*;*

i.e. the decision maker obtains a higher payoff if she follows her prior beliefs than if she follows the signals of an expert with reputation α. Our assumption also implies that an expert with a reputation of α will believe that state 0 is more likely regardless of his signal. This can create incentives for the expert to lie about his signal, thus impeding the decision maker's learning about his quality.

The timing of the interaction in each period is as follows. First, the decision maker decides whether to hire the expert. If he is employed, the expert then observes a signal and sends a subsequent cheap-talk report to the decision maker, after which the decision maker chooses a policy. Then, at the end of the period, the actual state of the world is publicly observed, and payoffs are realized.

Our solution concept is perfect Bayesian equilibrium. In addition,

Assumption 3.2 *We restrict attention to those equilibria in which the decision maker terminates the expert whenever the continuation value of continuing to employ him is 0.*

This restriction could be viewed as a shortcut that avoids the introduction of a small opportunity cost $c > 0$ of employing the expert. In some applications, this cost could e.g. represent exogenously specified wages, opportunity costs of the decision maker's time spent with the expert, or resources required to provide the expert with access to information. In some other applications, this restriction could be a consequence of external political pressures that make it impossible to retain an advisor who has proved himself to be incompetent. Indeed, without Assumption 3.2, the expert's career concerns would have no impact in our model because it would be optimal for the decision maker to grant the expert tenure in the first period.

In our environment, the decision maker faces two objectives. On the one hand, she chooses an optimal policy in each period given the available information. On the other hand, she chooses her employment strategy with a view toward minimizing the effect the expert's career concerns might have on his reports.

Achieving the first objective is straightforward and will not be the focus of our analysis: If the expert is employed, the decision maker will follow his recommendation if and only if it is sufficiently informative.

In particular, if the decision maker believes the expert is telling the truth, following his report is strictly optimal if and only if the decision maker thinks the signal is informative enough to overcome her prior, i.e.

$$\beta_t > 1 - p. \tag{1}$$

(Observe that Assumption 1 states that (1) does not hold in the first period.) If the expert's report is not sufficiently informative or if the expert is not consulted, the decision maker will follow her prior and choose policy 0.

4 First Best

Consider a hypothetical environment in which the expert's signals are observed by the decision maker. Let $\alpha_N(k)$ denote the posterior belief that the expert is competent at the beginning of the last period if there were k incorrect signals in the preceding periods. The value of $\alpha_N(k)$ is positive and decreasing in k if $q < 1$ and is equal 0 for any $k \geq 1$ if $q = 1$. The expert's signal in the last period is valuable for the decision maker if following the signal generates a higher expected payoff than following her prior

$$\alpha_N(k)q + (1 - \alpha_N(k))\frac{1}{2} > 1 - p. \tag{2}$$

To avoid uninteresting cases, we make

Assumption 4.1 *The inequality (2) is satisfied for $k = 0$.*

Definition Let κ be the highest $k \in \mathbb{N} \cup \{0\}$ for which (2) is satisifed.

Thus, κ is the maximal number of mistakes after which the expert's signal is valuable for the decision maker in the last period.

Definition The first-best decision rule

1. employs the expert until his reports have disagreed with the state $\kappa + 1$ times;
2. implements a policy equal to the expert's report if $\alpha_t > 1 - 2p$ and policy 0 otherwise.

If the expert's reports are truthful, this rule is a best response for the decision maker because it maximizes the decision maker's payoff and retains the expert if and only if the

decision maker's continuation value of doing so is positive. The first-best decision rule provides a natural benchmark against which we can assess the effect of the expert's career concerns. Furthermore, the decision maker's payoff if she follows the first-best decision rule and the expert reports his signals truthfully is the upper bound on her payoff in our model as well as in a richer model in which consulting an expert entails an opportunity cost (cf. our remarks after Assumption 3.2).

Definition The first-best decision rule is incentive compatible if there exists an equilibrium in which the decision maker follows this rule and the expert's reports are truthful for every history on the equilibrium path.

The agency problem in our model arises because the first-best decision rule might not be incentive compatible. Let, for instance, $N = 2$ and imagine that the expert observes $\tilde{s}_1 = 1$ in the first period. By Assumption 3.1, condition (1) is violated with slackness for $t = 1$ and, therefore, the expert believes that the state $\omega_0 = 0$ is more likely. Thus, the probability of employment in the next period is maximized by reporting $\hat{s}_1 = 0$. As a result, the expert's best response to the first-best decision rule would entail a report of 0 in period 1 irrespective of the observed signal.

Nevertheless, if a competent expert never makes mistakes,[6] the following proposition shows that if N exceeds a certain threshold, the first-best decision rule becomes incentive compatible. The result relies on what we call the *good news* effect: If the expert lies and his report turns out to be correct, he privately learns that he is incompetent. By contrast, if he reports his signal truthfully and it is correct, then the expert believes that he is more likely to be competent. Moreover, if the report is incorrect, the expert is fired and his continuation payoff is 0 regardless of his beliefs. While deciding about his report, the expert is concerned about about how likely he is to *continue* to be correct in the future if retained and this expectation is, therefore, conditioned on two different events. Thus, when there is "a lot of future left," i.e. in the early periods of a long game, this effect makes lying so unattractive as to obviate all incentive concerns.

Although there is an obvious analogy, the proof is not a folk-theorem type of argument. First of all, there is no discounting in our environment and the number of periods is finite. More importantly, the incentive problem disappears because of the different rates of growth

[6]This assumption implies that the decision maker will learn that the expert is certainly incompetent whenever he is expected to tell the truth and his report is inconsistent with the realized state. Similar full-revelation assumptions are commonly made, e.g. in the strategic experimentation literature, where it is often assumed that one instance of good (or bad) news resolves all uncertainty about the value of the risky arms (see e.g. Keller, Rady, Cripps, 2005, or Klein & Rady, 2010).

in the payoffs from lying and from telling the truth as the number of periods increases. The reason for this is that the expert evaluates his future payoffs conditional on different events.

Proposition 4.2 (Vanishing Career Concerns) *Assume that the competent expert never makes mistakes. For any given p and α, there exists an integer N_0 such that the first-best decision rule is incentive compatible if and only if $N \geq N_0$.*

PROOF: A formal version of the argument expounded above proves that for any t there exists an integer $N'(t)$ such that for all $N \geq N'(t)$ there is no profitable (possibly, multi-period) deviation from truth-telling that starts in period t. It is left to show, then, that there exists an N_0 such that $N(t) \leq N_0$ for all t or, in other words, that as we increase N the incentive constraints are not violated in the newly added periods. This, however, holds true because, if the expert is employed toward the end of the relationship under the first-best decision rule, then his reputation is necessarily high, the expert considers his signals very informative, and truth-telling is his strict best response. A complete proof is provided in the appendix. ∎

The insight that a longer time horizon solves the incentive problem is valid if the competent expert is always correct. However, if the competent expert might occasionally observe incorrect signals, this is no longer the case, as the following example shows. Here, the first-best outcome can be attained in equilibrium if $N = 2$ but not if $N = 3$.

Example Let $\alpha = 5/12$, $p = 3/7$, and $q = 9/10$.

1. Let $N = 2$. The first-best decision rule retains the expert in period 2 if and only if his signal is correct in period 1. This rule is incentive compatible.

2. Let $N = 3$. The first-best decision rule always retains the expert in period 2, and retains 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. In particular, if the decision maker follows this rule, the expert's best response after an incorrect signal in period 1 is to disregard his signal and report 0 in period 2.

In this example, the decision maker would like to continue to employ the expert if he makes a mistake in period 1 if $N = 3$ but not if $N = 2$. This is so because with more remaining periods there is a chance that the expert will prove himself to be sufficiently competent to become valuable for the decision maker. However, after a mistake, the expert is no longer willing to report his signal truthfully. If $N = 2$, this does not matter as the

expert is fired but if $N = 3$ the first-best decision rule ceases to be incentive compatible. This difficulty does not arise if $q = 1$, as then a single mistake fully reveals that the expert is of no value to the decision maker.

Our next proposition establishes that this problem will always arise if the number of periods is sufficiently large. The basic intuition is that the support of posterior beliefs in the final periods of the game becomes dense on $[0, 1]$ as the number of periods increases. Therefore, there will always exist a history after which the expert faces the same incentive problem as in our example with $N = 3$.

Proposition 4.3 (Persistent Career Concerns) *Suppose the competent expert occasionally makes mistakes, i.e. $q < 1$. Then, for any given p and α and any integer N_0, there exists $N \geq N_0$ such that the first-best decision rule is* not *incentive compatible.*

PROOF: See Appendix. ∎

This result does not address how much of a loss the decision maker will incur as a result of the incentive problem that precludes the first-best outcome. Although an incentive problem arises with positive probability after some histories, the set of these histories might become small as $N \to \infty$, and the decision maker might be able to remedy the problem by slightly modifying the decision rule.

Unfortunately, there is no simple solution. Recall that in the environment in which a competent expert never makes mistakes, the positive result is due to the good news effect. This effect implies that, in the first-best decision rule, the expert's expected continuation payoff from telling the truth diverged in N, while his payoff from lying stayed bounded. This is true for any belief about the expert's competence. However, this is no longer true when a competent expert may also make mistakes. Now, the continuation payoff from telling the truth at history nodes at which the expert is fired after one (or say k) additional mistake(s) does not diverge in N. As a result, there exists an α^* such that truth-telling is incentive compatible if and only if the expert's belief about his competence is higher than or equal to α^*. The difficulty is that the probability of reaching a belief less than α^* is fully determined by the prior belief; this probability does not vanish as $N \to \infty$.

This concludes the section on the first-best decision rules. We study the second-best decision rules in the next section.

5 Second Best

In this section, we assume a competent expert never makes mistakes.[7]

5.1 Fully Revealing Equilibria

However, even if the first-best decision rule is not incentive compatible, the decision maker could still try and enforce truth-telling by threatening to fire the expert with some probability after he has correctly forecast the more likely state. This threat is in turn made credible by the expert's off-path threat to babble and send but uninformative messages if he continues to be employed and the decision maker does not fire him when she is supposed to do so.[8]

In order to ensure that a deviation by the decision maker from firing the expert with the probability specified in equilibrium is observable, the game can be modified to allow for jointly controlled lotteries (Aumann and Maschler, 1995): We assume that after the decision maker takes an action but before the state is publicly observed, both players simultaneously send messages from the set $[0, 1]$ and these messages immediately become common knowledge.

Now imagine, for instance, that the decision maker is supposed to fire the expert with a probability of one half if his report is incorrect. In equilibrium, both players randomize between two messages with equal probability , e.g., 0 and 1. If the messages coincide, which happens with probability one half, the decision maker continues to employ the expert even after an incorrect report. Otherwise, the expert should be fired. These strategies achieve the desired probability of firing and, moreover, a deviation by the decision maker to any strategy that employs the expert with a different probability is publicly observable. Furthermore, a deviation to a different message strategy by a single player does not affect the expert's employment probability, and hence the reporting strategies are mutually best responses. A similar construction is possible for any firing probability, although it will typically require more complex message strategies.

Definition An equilibrium is fully revealing if on the equilibrium path the expert reports his signals truthfully.

The following result shows that truthful equilibria can indeed be constructed.

[7] This version of the paper is preliminary and we are currently working on the general case.

[8] This is incentive compatible for the expert, because the decision maker does not condition his continued employment on the reports.

Proposition 5.1 *There exists a fully revealing equilibrium in which on the equilibrium path the expert is always retained after a correct report* 0 *and retained with positive probability after a correct report* 1.

PROOF: We construct a truthtelling equilibrium with the following strategies.

Messages. The expert reports the true state unless he has observed a deviation by the decision maker, in which case he reports 0 in each period he is consulted. Let A_1 and A_2 respectively denote the messages sent by the expert and the decision maker in the jointly controlled lottery. Both players draw their messages A_1 and A_2 from the uniform distribution on $[0, 1]$.

Employment. On the equilibrium path, the decision maker fires the expert whenever the report is incorrect and retains the expert after a correct report 1. After a correct report 0, the expert is retained if and only if $A_1 - A_2 \in [0, \rho^s_{0,t}]$ or $A_2 - A_1 \in [1 - \rho^s_{0,t}, 1]$, where $\rho^s_{0,t}$ is constructed by induction in descending order of periods: In each period set its value equal to the maximal value with which the expert can be retained after correct report 0 such that reporting truthfully in this and future periods is incentive compatible given continuation strategies of the players. From the explicit expression for the incentive constraint, it is immediate that $\rho^s_{0,t}$ is well defined and is greater than $\frac{p}{1-p}$.

If the decision maker has deviated from her employment strategy, she never consults the expert.

Mutual best response property. First, consider a deviation by the decision maker from her employment strategy. This deviation is commonly known to trigger uninformative reports on the expert's part. This makes it optimal for the decision maker not to employ the expert in the future. In turn, even if the expert is consulted, reporting 0 is a best response as he expects not to be consulted ever again.

Second, neither player can affect his expected probability that the expert will be retained after a correct report of 0 by unilaterally deviating to a different distribution over his respective message A_i. Hence, these strategies are mutually best responses.

Next, by construction of message strategies, the expert is retained after a correct report 0 with probability $\rho^s_{0,t}$. Again, by construction, the probability is chosen such that truthful reporting is a best response for the expert.

Moreover, employing an expert who has not made an incorrect report is a best response for the decision maker by Assumption 4.1. ∎

However, providing incentives for the expert to report his information truthfully entails

a non-negligible cost for the decision maker. In the following proposition, we show that if $\kappa = 0$ and the first-best decision rule is not incentive compatible, the decision maker's payoff in any fully revealing equilibrium is bounded away from v^{FB}.

Let t^{FB} be the lowest number such that an expert who has observed t^{FB} correct signals will henceforth find it optimal to report his signals truthfully if the decision maker follows the first-best decision policy. Clearly, if $t^{FB} = 0$, the first-best decision rule is incentive compatible. Furthermore, $t^{FB} \leq N - 1$ because the expert is indifferent about his report in the last period.

The proof of the proposition relies on the fact that if $t^{FB} > 0$, there necessarily exists a history along the path of play after which the expert is fired with positive probability even though he has not made a mistake. The upper bound we give then follows from a situation where this history is unique.

Proposition 5.2 *Suppose $\kappa = 0$. Let $t^{FB} > 0$, and let V, and ρ, be the decision maker's payoff and strategy, respectively, in a fully revealing equilibrium. Then, there exists a time $t < N$, a constant μ_t and a history $h_t \in \mathcal{H}_1$ such that*

$$V \leq v^{FB}(\alpha, N) - \mu_t(1 - \rho(h_t))v^{FB}(\alpha_{t+1}, N - t) < v^{FB}(\alpha, N).$$

PROOF: Since $t^{FB} > 0$, there exists some $t < N$ and some history $h_t \in \mathcal{H}_1$ with $\rho(h_t) < 1$ and $\mu_t > 0$, where μ_t denotes the probability of history h_t realizing in this equilibrium. The upper bound is then given by the hypothetical situation where h_t is the only history after which the first-best policy is distorted. By assumption 4.1, $v^{FB}(\alpha_{t+1}, N - t) > 0$; hence V is bounded away from $v^{FB}(\alpha, N)$. ∎

Thus, the decision maker's payoff in any fully revealing equilibrium will be bounded away from v^{FB}. In the next subsection, we shall construct an equilibrium which entails the accumulation of endogenous private information on the agent's part. We show that this equilibrium attains the payoff of v^{FB} for the decision maker and hence is the best possible equilibrium.

5.2 Equilibrium With Endogenous Private Information

A quite natural way for the decision maker to handle the expert's incentive problem would be for her to grant him an initial "grace period", during which he was allowed to send uninformative signals each period, and to gain confidence in his abilities, finding his mark

in his new job. Once this probationary period, which lasts for t^{FB} periods, ends, though, he is expected to be right every time, i.e. the first-best decision rule will be implemented. The expert will then report his signals truthfully if, and only if, his signals have all been correct during the first t^{FB} periods; otherwise, he will best respond by continuing to babble, i.e. to announce state 0 no matter what his signal may have been.

We summarize this equilibrium in the following proposition:

Proposition 5.3 *There exists an equilibrium in which no information is transmitted, and the expert is never fired during the first t^{FB} periods; thereafter, the expert truthfully reveals his signals if, and only if, his first t^{FB} signals were correct. Moreover, he will only be fired as soon as he has made an incorrect forecast after the first t^{FB} periods.*

PROOF: The expert's equilibrium strategy is for him always to report state 0, unless $t > t^{FB}$, and he has been employed, and has received correct signals, in all previous periods, in which case he truthfully reveals his signal. The decision maker's equilibrium strategy calls for employing the expert in $t = 1$, and for retaining him if he reports 0 during the first t^{FB} periods. In all other cases, she retains the expert if, and only if, his current prediction turns out to be correct. These strategies are mutually best responses by the definition of t^{FB} and the fact that the decision maker learns nothing about the competence of the expert in all periods $t \leq t^{FB}$. ∎

The decision maker's policy choices in this equilibrium are those she would make in the first-best environment: Suppose the expert has privately learned that he is of the bad type; he then maximizes his expected employment duration by always reporting state 0. A decision maker who knew that she could not rely on the expert's advice would also optimally stick with her prior and implement policy 0. By contrast, in any fully revealing equilibrium, a good expert will be fired with positive probability, leading to worse policy decisions in expectation from an ex-ante point of view. Thus, on account of the agency problem, the ex-ante expected duration of employment is longer than in the first best.

Indeed, it can only be to the principal's advantage for the agent to be better informed, even if this information be held privately; an expert who is more optimistic will be more inclined to reveal his signal, and following his signal is a good idea for the principal also. A privately pessimistic expert by contrast will tend to follow his prior without any regard to his signal; in this case, following her prior belief is also the best the principal can do in terms of policy. If, on the other hand, the principal's goal were to screen out a bad expert, private information would rather tend to hurt the principal.

Thus, even though the first-best decision rule may not be incentive compatible, this equilibrium still achieves the first-best payoff for the decision maker. Still, it violates condition 1. of our definition of the first best; indeed, it is second-best because the expert is employed longer in expectation than in the first-best decision rule (recall our discussion after Assumption 3.2: we made the point that our model could be viewed as a reduced-form model of an environment in which consulting an expert entails a small cost for the decision maker). Of course, if the decision maker incurred such a (small) cost for employing the expert, she would prefer firing a bad expert as quickly as possible. As it turns out, it is impossible to achieve v^{FB} while employing the expert for fewer expected periods than is the case in our equilibrium, as the following corollary shows. Thus, this equilibrium would continue to be second-best in a richer model with employment costs provided these costs were sufficiently small.

Corollary 5.4 *If $\kappa = 0$, the decision maker's ex-ante expected payoff in the equilibrium identified in Proposition 5.3 is equal to v^{FB}. Furthermore, there does not exist an equilibrium in which the decision maker obtains the same ex-ante expected payoff and the ex-ante probability of retaining the expert is smaller.*

PROOF: The first statement immediately follows from our previous discussion. Regarding the second statement, suppose on the contrary that there exists an equilibrium achieving v^{FB} in which the expert is employed for fewer periods in expectation. In order for the principal to achieve an ex-ante expected value of v^{FB}, it must be the case that a good expert is never fired; i.e. in such an equilibrium, the expert is only fired after he has revealed himself to be of the bad type. Since he is employed for fewer periods in expectation than in the equilibrium exhibited in Proposition 5.3, it must be the case that some information on the agent's type will be transmitted in period t^{FB} or earlier. Yet, to induce the expert to tell the truth with some positive probability in period t^{FB} or earlier, the decision maker has to fire the expert with some positive probability even after he has been correct. This in turn implies that a good expert will be fired with positive probability, and that hence the decision maker's payoff in this equilibrium is bounded above by

$$v^{FB}(\alpha, N) - \mu_t(1 - \rho(h_t))v^{FB}(\alpha_{t+1}, N - t - 1),$$

for some $\mu_t, \rho(h_t) > 0$. ∎

6 Conclusion

We have investigated the dynamic interaction between a decision maker and an expert of unknown quality who privately observes a potentially decision-relevant signal. As he only cares about his reputation insofar as it translates into a longer expected duration of employment, the expert may have an incentive strategically to manipulate his cheap-talk relay of the signal to the decision maker. We have shown that if a competent expert never makes mistakes and the number of periods is large enough, the expert's career concerns vanish, and the first-best becomes implementable. For short time horizons, the decision maker may be better off letting the expert accumulate some private information during an initial grace period than he would be in any fully revealing equilibrium.

Appendix

Proof of Proposition 4.2

Suppose the decision maker pursues the first-best policy of immediately firing the expert if, and only if, the expert has made a mistake. Then, the agent is willing to reveal a signal indicating the less likely state 0 truthfully at any time t, if at all times $0 \leq t \leq N$, the following incentive constraint holds:

$$p\left[\alpha_t(N-t) + \frac{1-\alpha_t}{2}\left(1 + \tfrac{1}{2} + \cdots + \frac{1}{2^{N-t-1}}\right)\right] \geq (1-p)\frac{1-\alpha_t}{2}\left[1 + (1-p) + \cdots + (1-p)^{N-t-1}\right]. \quad \text{(A.1)}$$

To understand the right-hand side of the incentive constraint, the reader should note that if, upon lying, the expert finds out *ex post* that his message was in fact correct, he then privately learns that he is of the low type and will maximize his continuation payoff by reporting the *a priori* more likely state in all subsequent periods.

It is now immediate to verify that, as $N \to \infty$, the left-hand side diverges to $+\infty$, whereas the right-hand side converges to $\frac{1-p}{p}\frac{1-\alpha_t}{2} < \infty$. Let N_0 be the smallest value of N for which this constraint is satisfied for all $t \leq K$. By our assumption that $\beta_0 < 1-p$, we have $N_0 \geq 2$.

It is left to check that the constraint is also satisfied for all $t > K$. It is direct to verify that the constraint holds for any N if $\alpha_t = 1-2p$. Furthermore, the left hand side of the constraint is increasing in α_t while the right hand side is decreasing in α_t. Therefore, the constraint is satisfied for all $\alpha_t \geq 1-2p$, which is equivalent to $t \geq K$.

As is straightforward to verify, the left-hand side of the incentive constraint conditional on a signal indicating the more likely state 1, is $\frac{1-p}{p} > 1$ times the left-hand side of the above constraint, whereas the right-hand side is $\frac{p}{1-p}$ times the above right-hand side. Therefore, this constraint also holds for all $N \geq N_0$. ∎

Proof of Proposition 4.3

Assume that the decision maker follows the first-decision rule and that the expert reports his information truthfully. Consider a history at the beginning of period $N-1$ with k correct signals and $N-2-k$ incorrect signals. Then,

$$\hat{\beta}_{N-1-m} = \beta_{N-1-m} = \frac{\alpha q^{k+1}(1-q)^{N-2-m-k} + (1-\alpha)\frac{1}{2^{N-1-m}}}{\alpha q^{k}(1-q)^{N-2-m-k} + (1-\alpha)\frac{1}{2^{N-2-m}}}$$

If the expert makes one more correct report in period $N-1$, we have

$$\hat{\beta}_{N} = \beta_{N} = \frac{\alpha q^{k+2}(1-q)^{N-2-k} + (1-\alpha)\frac{1}{2^{N}}}{\alpha q^{k+1}(1-q)^{N-2-k} + (1-\alpha)\frac{1}{2^{N-1}}}$$

The first-best decision rule is not incentive compatible if (1) is violated with a strict sign for

$\beta_t = \beta_{N-1}$ but holds for $\beta_t = \beta_N$ or, equivalently, if

$$\left(\frac{q}{1-q}\right)^k < \frac{1-p-\frac{1}{2}}{q-(1-p)}\frac{1-\alpha}{\alpha} \cdot \frac{1}{(2(1-q))^{N-2}} < 2q \cdot \left(\frac{q}{1-q}\right)^k.$$

which reduces to

$$0 < z_1 + (N-2)z_2 - k < z_3 \tag{A.2}$$

where $z_1 = \log_{\frac{q}{1-q}} \left[\frac{1-p-\frac{1}{2}}{q-(1-p)}\frac{1-\alpha}{\alpha}\right]$, $z_2 = -\log_{\frac{q}{1-q}} 2(1-q)$, and $z_3 = \log_{\frac{q}{1-q}} 2q$. We have $z_1 > 0$ by Assumption 3.1, and $0 < z_3 < z_2 < 1$ and $z_2 + z_3 > 1$ as $q > (1-p) > \frac{1}{2}$.

Now, all that remains to be shown is that there exists a pair $(N,k) \in \mathbb{N}_0^2$ such that both inequalities in (A.2) hold. This is easy to show, however. First suppose that $z_1 < z_3$. Then, $(N,k) = (2,0)$ does the trick. If, however, $z_1 \geq z_3$, we choose $n \in \mathbb{N}$ such that $0 < z_1 - n(1-z_2) < z_3$. (As $1 - z_2 < z_3$, such an $n \in \mathbb{N}$ exists.) Now, setting $(N,k) = (n+2, n)$ completes the proof. ∎

References

Bar-Isaac, H. (2003): "Reputation and Survival: Learning in a Dynamic Signalling Model," *Review of Economic Studies*, 70, 231–251.

Bergemann, D. and U. Hege (2005): "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36, 719–752.

Bergemann, D. and U. Hege (1998): "Dynamic Venture Capital Financing, Learning and Moral Hazard," *Journal of Banking and Finance*, 22, 703–735.

Bergemann, D. and J. Välimäki (2008): "Bandit Problems," in: *The New Palgrave Dictionary of Economics*, 2nd edition, ed. by S. Durlauf and L. Blume. Basingstoke and New York, Palgrave Macmillan Ltd.

Bergemann, D. and J. Välimäki (2000): "Experimentation in Markets," *Review of Economic Studies*, 67, 213–234.

Bergemann, D. and J. Välimäki (1996): "Learning And Strategic Pricing," *Econometrica*, 64, 1125–1149.

Bolton, P. and C. Harris (2000): "Strategic Experimentation: the Undiscounted Case," in: *Incentives, Organizations and Public Economics – Papers in Honour of Sir James Mirrlees*, ed. by P.J. Hammond and G.D. Myles. Oxford: Oxford University Press, 53–68.

Bolton, P. and C. Harris (1999): "Strategic Experimentation," *Econometrica*, 67, 349–374.

Bradt, R., S. Johnson and S. Karlin (1956): "On Sequential Designs for Maximizing the Sum of n Observations," *The Annals of Mathematical Statistics*, 27, 1060–1074.

Gerardi, D. and L. Maestri (2008): "A Principal-Agent Model of Sequential Testing," Cowles Foundation Discussion Paper No. 1680.

Hörner, J. and L. Samuelson (2009): "Incentives for Experimenting Agents," Cowles Foundation Discussion Paper No. 1726.

Keller, G. and S. Rady (2010): "Strategic Experimentation with Poisson Bandits," *Theoretical Economics*, 5, 275–311.

Keller, G., S. Rady and M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73, 39–68.

Klein, N. and S. Rady (2010): "Negatively Correlated Bandits," *Review of Economic*

Studies, forthcoming.

Klein, N. (2010): "Strategic Learning in Teams," working paper, available at http://ssrn.com/abstract=

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, 231–265.

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

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

Rothschild, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185–202.