Too good to be truthful: Why competent advisers are fired*

Christoph Schottmüller†
University of Cologne

February 28, 2019

Abstract

A decision maker repeatedly asks an adviser for advice. The adviser is either competent or incompetent and knows his type privately. His preferences are not perfectly aligned with the decision maker's preferences. Over time, the decision maker learns about the adviser's type and will fire him if the adviser is likely to be incompetent. If the adviser's reputation for competence improves, he is less likely to be fired for incompetence but this makes pushing his own agenda more attractive to him. Consequently, very competent advisers are also fired with positive probability because they are tempted to pursue their own goals. The quality of advice can be highest if the adviser's competence is uncertain.

JEL codes: C73, D83, G24

Keywords: advice, cheap talk, reputation

1. Introduction

As specialization is one of the cornerstones of the modern knowledge society, it is not surprising that advice provided by specialized experts is important in so many domains of life. Savers have financial advisers to help them manage their wealth, consumers turn to sales personnel to make good choices, politicians and managers depend on their advisers to find the right policies, patients need their physicians' advice and internet users rely on search engines.

In most instances, the adviser's incentives are not necessarily aligned with the advice seeker's preferences. Financial advisers, sales personnel and search engine operators can

*I want to thank Ole Jann, Jesper Rüdiger, Peter Norman Sørensen, Juuso Välimäki, audiences at the University of Copenhagen, SING12 in Odense and the Meeting of the EEA in Geneva, two anonymous referees and the associate editor for helpful comments. Support by the Danish Research Council (grant number: DFF–7015-00020) is gratefully acknowledged.

†email: c.schottmueller@uni-koeln.de, phone: +49 22 14 70 60 68

win bonuses if their customers buy certain products, while politicians and managers might wonder whether their advisers have an agenda of their own and patients might be worried that their physician's enthusiasm for a certain drug stems from successful lobbying efforts on the part of its producer. Even ex post, it is hard to detect whether such concerns are justified, because advice in these contexts is complex and even the best possible advice can turn out to be wrong once in a while.

Another common feature of these examples is the repeated nature of the advice. Most people tend to receive advice from the same adviser several times and switch advisers only from time to time. It clearly makes sense to switch advisers if the advice seeker concludes that his adviser is not competent. However, long term advisers who are viewed as very competent are also occasionally fired. For example, in 2003, financial analyst Jack Grubman was banned by the Security and Exchange Commission from the financial industry for life and fined $15 million for misconduct. Grubman had used his good reputation to pursue his personal goals when he gave a public buy recommendation for AT&T as part of a complex plan to enable his children to be admitted to the prestigious 92nd Street YM-YWHA's preschool program (as he explained in a private email that later went public).[1] By the time the ban was announced, market participants had, of course, already stopped listening to Grubman's advice but this reaction was not a response to perceived incompetence. When Grubman was hired by Distinctive Devices as consultant a year later, the company's stock price went up. The problem was that Grubman had apparently (ab-)used his good reputation: he had misrepresented his information and thereby had manipulated his followers for his own personal benefit.

History provides an ample supply of examples of kings who dismissed or even killed their most prominent advisers when those advisers were too competent and perceived as a threat to the throne. Most infamous in this respect was the Ottoman Sultan Suleiman the Magnificent, who killed not only his Grand Vizier and childhood friend Pargali Ibrahim Pasha but also his own son and designated heir Mustafa for this reason (after a successful military campaign on his father's behalf during which Mustafa committed the mistake of not stopping his soldiers from referring to him as "Sultan").

In these examples, a competent adviser was mistrusted and fired after engaging in activities that made the decision maker doubt whether the adviser acted in the decision maker's best interest or whether he was instead (ab-)using his power to push his own agenda. This paper argues that instances such as these are typical. More specifically, advisers can be fired not *in spite of* having a reputation for being competent but *because* of it. That is, they might have given better advice and kept their positions if their competence had been in doubt.

[1]See `http://observer.com/2010/03/stockgoosing-grubman-to-sell-townhouse-for-196-m/` for a brief summary of the story. (Accessed on Feb. 20, 2019)

What is the logic behind this result? In this paper, I consider a setting in which the competence of an adviser is not perfectly known by the decision maker. An adviser whose competence is in doubt faces the danger of being dismissed for incompetence if his advice turns out to be bad (thus reinforcing the decision maker's initial doubts). Therefore, the adviser has a strong incentive to act in the decision maker's best interest in order to keep his position. An adviser who is believed – with high probability – to be competent, however, has more latitude because the risk of him being fired *due to incompetence* in the near future is negligible. In other words, even if his advice turns out to be poor a few times, this is not immediately a sign of incompetence as it could simply be due to bad luck. The adviser is therefore free to pursue his own goals, which are usually not in line with the decision maker's goals. But in this case, the best response of the decision maker is to fire the adviser whose advice serves principally his own interests rather than those of the decision maker.[2]

A competent adviser can therefore lose his job for two distinct reasons. If the belief that the adviser is competent is too low, the decision maker will fire the adviser because the information that the adviser is competent is hidden. If the belief that the decision maker is competent is high, the decision maker will fire the adviser not because of hidden information but because of moral hazard: the adviser does not act in the interests of the decision maker but pushes his own agenda.

The model is a repeated game in which the adviser recommends one of two options to the decision maker in every period until the decision maker ends the advice relationship. One of the available options fits the decision maker's needs and one fits the adviser's needs – e.g. he receives a bonus when recommending that option. The two might accidentally coincide from time to time but often they do not. The decision maker has a uniform prior concerning which option fits his needs and also concerning which option fits the adviser's needs. The adviser receives a noisy signal indicating which option fits the decision maker's needs and he knows perfectly which option will give him a bonus. The decision maker discovers whether the recommended option has fitted his needs only after he has followed the recommendation. The adviser has one of two types: either he is competent – i.e. his noisy signal is informative – or not.

In this model, no meaningful advice can be obtained in a static setting because the adviser will always recommend his bonus option if he does not face the threat of losing future bonus payments. The same is true in a finitely repeated game: as in the static setting, the adviser is unable to give meaningful advice in the last period and as a

[2]To be more precise, the decision maker and adviser might in equilibrium use mixed strategies when the adviser is believed to be competent. Thus, the decision maker will fire the adviser with some probability when a recommendation turns out to be poor. This gives the adviser some incentive to not give too poor advice to avoid being fired. In equilibrium, the quality of advice will be just high enough to make the decision maker indifferent between firing and keeping the adviser. Otherwise, the decision maker's threat of firing would not be credible.

consequence, he will always be fired before the last period. Given this, the adviser is unable to give meaningful advice in the second to last period and the game unravels, meaning that the adviser is never consulted in equilibrium. Some informative advice is, however, possible in an infinitely repeated game setting. Unsurprisingly, the adviser is fired for sure if the decision maker's belief that the adviser is competent is very low. If this belief is sufficiently high, then the adviser is also fired with positive probability – at least in the event that he recommends an option that does not fit the decision maker's needs. For these high beliefs, equilibrium strategies are usually mixed. The decision maker is indifferent between firing and retaining the adviser and the threat of firing is just high enough to ensure that the adviser finds a strategy optimal that keeps the decision maker indifferent between these two options.

The expected length of the game, i.e. the number of periods before the adviser is fired, is uniformly bounded from above for any belief; that is, the bound is independent of the decision maker's belief regarding the adviser's competence. This illustrates that even an arbitrarily competent adviser will almost surely be fired within a finite amount of time. These results hold for all equilibria of the game, i.e. they are not affected by multiplicity of equilibria. The adviser suffers in many equilibria from a severe commitment problem: if he was able to commit to truthfully revealing his signal in every period, then he and the decision maker would both obtain strictly higher payoffs than in equilibrium.

The outline of the paper is as follows. The next section introduces the model. Section 3 presents the results holding in all perfect Bayesian equilibria and in Markov equilibria in particular (section 3.1). Section 3.2 discusses several generalizations that are pursued in greater detail in the supplementary material to this paper. The model is extended to allow for monetary transfers and competition among several advisers in section 4. Section 5 discusses the related literature and section 6 concludes. Proofs as well as some technical discussions are relegated to the appendix.

2. Model

Timing: In an initial period 0, the decision maker (DM) decides whether to consult the adviser (A) at all or to take the outside option immediately. If A is consulted, players play the stage game described below in each following period until DM ends the game by firing A. Figure 1 gives an overview of the stage game where payoffs are depicted in the top row, the players' actions in the middle and changes in information in the bottom row.

Stage game (actions): As long as DM does not fire A, the stage game in period t is as follows: Firstly, A receives a noisy signal regarding DM's needs in this period.

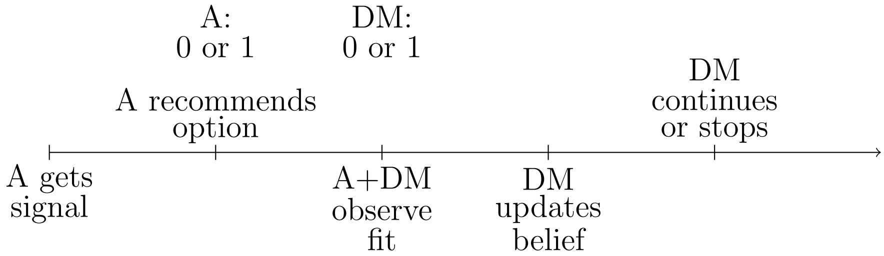

Figure 1: Timeline within a given period t.

Secondly, A recommends one of two available options to DM. Thirdly, DM decides whether to end the game or to continue to the next period, i.e. whether to fire the adviser or not.

Stage game (payoffs): Only one of the two options fits DM's needs and DM receives a payoff of 1 (0) if the recommended option fits (does not fit) his needs.[3] One of the two options leads to a "bonus" for A. In other words, A receives a payoff of 1 (0) if he recommends (does not recommend) the bonus option. If DM ends the game, he will receive an outside option payoff of W_O in $t+1$ and a zero payoff thereafter while A's payoff will be zero in all future periods. Both players discount future payoffs using the common discount factor $\delta \in (0,1)$.

Stage game (information): Each option fits DM's needs with probability 1/2 and is A's bonus option with probability 1/2. These events are independent. The identity of the bonus option is private information of A. At the start of the period neither player knows which option fits DM's needs.

A has one of two types. If A is *incompetent*, his signal regarding DM's needs is completely uninformative. An incompetent type's belief assigns therefore probability 1/2 to each option fitting DM's needs. If A is *competent*, his signal is noisy but informative. More precisely, a competent type's belief concerning DM's needs assigns probability $p \in (1/2, 1)$ to one option and $1-p$ to the other option. I will refer to the more likely option as "option h" ($h$igh probability fit) and the less likely option as "option l" in the remainder of this paper (but keep in mind that DM is not aware of these identities as A's signal is private!). For simplicity, p is assumed to be time invariant. A's type is private information of A and also time invariant.

Whether the recommended option fits DM's needs is observed by both players di-

[3]The recommendation is directly payoff relevant. This is reasonable especially when the advice reveals the existence/identity of the option and DM is unaware of the identity of the not recommended option, see also footnote 4. The supplementary material explores how the setup can be reinterpreted as a cheap talk setting where DM has a (pseudo-) decision to choose an option after receiving the recommendation and this decision is payoff relevant.

rectly after A's recommendation when DM's payoff is realized. The identity of the bonus option, however, remains private information of A. Note that – due to the independence assumption – DM cannot infer from the identity of the recommended option the likelihood that A recommended the bonus option.[4] DM's needs and A's bonus options are uncorrelated across periods.

DM's belief as to whether A is competent is denoted by $\alpha \in [0, 1]$. This belief is updated using Bayes' rule after observing whether or not the recommendation fitted DM's needs. I will denote by α^+ (α^-) the updated belief after a recommendation fitted (did not fit) DM's needs. DM's initial belief α_0 is assumed to be common knowledge. Consequently, α will be commonly known in equilibrium as both players observe whether the recommendation fitted DM's needs. I will occasionally refer to α as *A's reputation.*

Parameter assumption: To make the problem interesting, DM's outside option is assumed to satisfy the inequalities

$$\frac{1/2}{1-\delta} < W_O < \frac{p}{1-\delta}. \tag{1}$$

That is, DM prefers his outside option to choosing an option randomly in each period, but he prefers choosing an option fitting his needs with probability p in each period to his outside option. Consequently, DM would like to fire an incompetent adviser, but would prefer to keep a competent adviser if that adviser reveals his signal truthfully in every period.

3. Analysis

I use perfect Bayesian equilibrium as main solution concept. In any such equilibrium the incompetent type always recommends his bonus option: the incompetent type cannot influence the probability with which DM's needs are satisfied because both options are equally likely to do so from his point of view. Consequently, recommending his bonus option yields – in expectation – the same consequences for future play as recommending the other option, but the bonus option gives an immediate payoff of 1.

Lemma 1. *In every perfect Bayesian equilibrium, the incompetent type recommends the bonus option in every period (after every history).*

[4] One interpretation is that DM only knows that there are 2 options but is unaware of what these options are; that is, he learns that a specific option exists only if it is recommended to him (this is the reason why he seeks advice in the first place). Think of a person googling "Italian restaurants in Manhattan" or a patient asking a physician for the right medication. Taking a non-suggested option is simply not feasible because it is unknown to DM in these examples. Advice can therefore be viewed as an experience good.

As the incompetent type's behavior does not require any further analysis I will refer with "adviser" (A) to the competent type in the remainder of the analysis.

For ease of exposition, I will restrict myself to "informative" perfect Bayesian equilibria which I define as perfect Bayesian equilibria in which a competent adviser gives at least as good advice as an incompetent one after any history. In other words, the competent type is at least as likely to recommend the option fitting DM's needs as the incompetent type.[5] I will refer to informative perfect Bayesian equilibria as "equilibrium" from here onward.

If the adviser's reputation is sufficiently low, it is optimal for DM to stop the game. This follows almost directly from the assumption that DM's outside option is strictly better than receiving advice from an incompetent adviser forever, see (1). If DM is almost sure to face an incompetent adviser, firing the adviser will therefore be optimal for him.

Proposition 1. *In equilibrium, there exists an* $\underline{\alpha} > 0$ *such that DM ends the game whenever* $\alpha < \underline{\alpha}$.

The more interesting result is that DM will also end the game (with some probability) if the adviser's reputation is sufficiently high. The intuition for this result is as follows: suppose DM continued for sure if α is above a certain threshold $\tilde{\alpha} < 1$. For α close enough to 1, A would then be very sure that DM will continue even if he gives bad advice repeatedly. This is true as α^- is very close to α if α is close to 1. Put differently, A has hardly any dynamic incentives to give good advice. Statically, however, he has an incentive to recommend the bonus option as this will give an immediate payoff of 1. A will therefore recommend the bonus option no matter what his signal is. Consequently, both types of adviser behave in the same way which has two implications: firstly, the belief updating stops, i.e. $\alpha = \alpha^+ = \alpha^-$, and secondly, DM's expected payoff is below his outside option, since this situation gives him the same payoff as receiving advice from an incompetent adviser forever. Clearly, this contradicts our starting point that DM continues whenever $\alpha > \tilde{\alpha}$. As such an $\tilde{\alpha} < 1$ does not exist, we can conclude that there are beliefs arbitrarily close to 1 where DM quits the game with positive probability. To state the result more formally define $\eta = (1-\delta)W_O - 1/2$ and note that $\eta \in (0, p - 1/2)$ by (1).

Theorem 1. *Let* $\bar{\alpha} \in (1/2, 1)$. *Then there does not exist an equilibrium in which the adviser is fired with a probability less than* $\varepsilon = (\eta - \eta\delta)/(2 - 2\eta\delta) > 0$ *after all histories after which his reputation* α *is greater than* $\bar{\alpha}$.

[5] While non-informative equilibria are somewhat nonsensical in economic terms, they can in principle exist because a competent A has the ability to give worse advice than an incompetent A. If DM expects A to do so (at some point of the game), then it might be a best response for A to give bad advice (at this point of the game) to improve (!) his reputation: fitting advice would then be interpreted as being more likely to be given by an incompetent type and therefore reduce α.

The idea behind the theorem and the specific threshold is that belief updating becomes arbitrarily slow if A's reputation is sufficiently close to 1. That is, for any $\bar{\alpha} < 1$ there are reputation levels α^* close enough to 1 such that α cannot drop below $\bar{\alpha}$ within T periods. If DM were to fire A with probability less than ε for reputation levels above $\bar{\alpha}$, then an adviser with reputation α^* could guarantee himself a payoff of $(1-\delta^{T+1})/(1-\delta(1-\varepsilon))$ by recommending his bonus option for the next T periods. The proof of theorem 1 shows that for $\varepsilon = (\eta - \eta\delta)/(2-2\eta\delta)$ this lower bound on the adviser's equilibrium payoff is so high that the value of continuing for DM at reputation level α^* must be less than his outside option as a high adviser payoff implies that the adviser recommends his bonus option even if it does not fit DM's needs. Hence, it is not optimal for DM to fire with probability less than ε at α^*.

The previous theorem established that DM will fire the adviser with a positive probability (bounded away from zero) for high α. However, if this probability was small and limited to only certain histories, one could argue that it has little economic relevance. The intuition given above should already illustrate that this is not the case. The following theorem strengthens this intuition by stating that DM quits the advice relationship almost certainly within T periods – where T is some finite number depending on the parameters – no matter what the current belief is. This result is then used to derive an upper bound on the expected length of the advice relationship. Note that these bounds depend neither on the (initial) belief α_0 nor on the equilibrium, i.e. they hold for every equilibrium and every initial belief.

Theorem 2. *Let $\varepsilon \in (0,1)$ and define*

$$T_\varepsilon = \left\lceil \frac{\log(\varepsilon)}{\log\left(1-(1-p)^{T'}\varepsilon'\right)} \right\rceil T' \qquad \textit{where} \quad \varepsilon' = \frac{\eta(1-\delta)}{2-2\delta\eta} \quad \textit{and} \quad T' = \left\lceil \frac{\log(\eta/2)}{\log(\delta)} \right\rceil .$$

The probability that DM ends the game within T_ε periods is at least $1-\varepsilon$ in every equilibrium.

The expected length of the advice relationship in equilibrium is finite and bounded from above by

$$\bar{T} = T'\left(2 - \frac{1}{\log(1-(1-p)^{T'}\varepsilon')}\right).$$

The previous theorem is based on the following idea: if it is sufficiently unlikely that A is fired in the event of repeatedly giving bad advice in the following T periods, then A finds it optimal to repeatedly recommend his bonus option even if it is option l. However, in this case DM should fire A immediately. This logic combined with the noisiness of A's signal yields an upper bound on the probability that the game continues in equilibrium for more than T periods (for sufficiently large T). This upper bound can then be used to compute the upper bound on the expected length of the game $\bar{T}$.

Theorems 1 and 2 establish inefficiencies that are unavoidable in equilibrium: For high levels of reputation, DM would like to continue if A could credibly commit to giving the best possible advice in the future and A would prefer this commitment to being fired. However, this is incompatible with equilibrium incentives. The results do, of course, not contradict the Folk theorem. If players are arbitrarily patient, i.e. $\delta \to 1$, the inefficiency bounds become arbitrarily slack. Specifically, ε in theorem 1 approaches zero and T' in theorem 2 approaches infinity as $\delta \to 1$.

Example 1. *I present here a family of simple equilibria and give a condition on the parameters under which those equilibria exist. On the equilibrium path, play is as follows: if $\alpha < \underline{\alpha}$, DM ends the game and A recommends his bonus option regardless of his signal.*[6] *If $\alpha \in [\underline{\alpha}, \bar{\alpha}]$, DM continues if the current period's recommendation fitted his needs and ends the game if it did not fit his needs. A recommends option h regardless of his signal. If $\alpha > \bar{\alpha}$, players use mixed strategies. More precisely, DM continues with probability*

$$\beta^{+} = \frac{1}{\delta(2p - 1/2)}$$

if the current period's recommendation fitted his needs and will end the game if it did not fit his needs. A recommends option h with probability 1 if it is the bonus option and with probability

$$s(\alpha, l) = \frac{(1-\delta)W_O - 1/2}{\alpha(p - 1/2)}.$$

if it is not. Detectable deviations, i.e. DM continuing after a non-fitting recommendation or after $\alpha < \underline{\alpha}$ in some earlier period, lead to the following off equilibrium path play: A will recommend his bonus option regardless of his signal and DM will end the game regardless of his belief.

Proposition 2. *If $p > (1+2/\delta)/4$, the strategy profile described above is an equilibrium for any $\underline{\alpha}$ and $\bar{\alpha}$ such that $((1-\delta)W_O - 1/2)/(p - 1/2) < \underline{\alpha} < \bar{\alpha} < 1$.*

In this equilibrium, DM ends the game whenever he receives a non-fitting recommendation (and – with some probability – even when receiving a fitting recommendation at high beliefs). Due to the noisiness of A's signal technology, this will happen eventually. DM prefers – on the equilibrium path – advice strictly to his outside option if $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ but is indifferent between equilibrium advice and his outside option if $\alpha > \bar{\alpha}$. That is, DM prefers an adviser with a mediocre reputation to an adviser with a high reputation. The quality of the advice that a competent type gives decreases with his reputation: Advice is worse for $\alpha > \bar{\alpha}$ than for $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ and the quality of advice, i.e. $s(\alpha, l)$, is

[6] Note that strategies are stated here in terms of the belief of a player at the time he has to act, i.e. DM's strategy depends on the updated belief after observing whether A's recommendation fitted his needs.

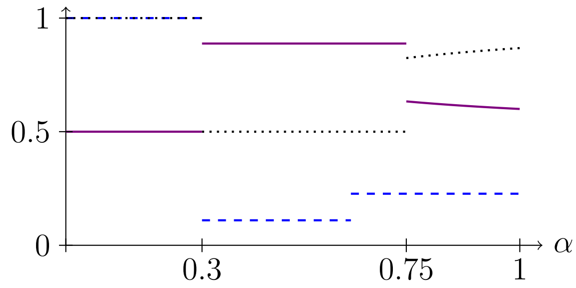

Figure 2: As a function of reputation *at the start of a period:* (i) probability that a competent type is fired this period (dashed blue line), (ii) probability of fitting advice by a competent adviser this period (solid, violet) and (iii) probability that a competent adviser recommends his bonus option this period (dotted, black). Parameter values: $\delta = 0.9$, $p = 8/9$, $W_O = 6$, $\underline{\alpha} = 0.3$ and $\bar{\alpha} = 0.75$.

strictly decreasing in α for $\alpha > \bar{\alpha}$. An adviser can be fired because *he is competent: for $\alpha > \bar{\alpha}$, the probability of being fired is higher than for $\alpha \le \bar{\alpha}$ due to two different effects. First, A gives worse advice and therefore dismissal due to bad advice is more likely. Second, DM uses a mixed strategy and might fire A even after fitting advice. Figure 2 illustrates these properties.*

The idea behind the construction of this equilibrium is simple. After histories in which either A gave unfitting advice in an earlier period or histories after which $\alpha < \underline{\alpha}$ a "babbling equilibrium" emerges, i.e. A always recommends his bonus option and DM always ends the game which are mutually best responses. For $\alpha > \bar{\alpha}$ a mixed strategy equilibrium is used where β^+ is just high enough to keep A indifferent between the two recommendations in case option l is the bonus option and $s(\alpha, l)$ is just high enough to keep DM indifferent between continuing and ending the game. The parameter assumption in proposition 2 ensures that these mixed strategies are feasible. For $\alpha \in [\underline{\alpha}, \bar{\alpha}]$, DM continues because a competent adviser will give best possible advice in the next period and the probability of facing a competent types is sufficiently large. A competent adviser on the other hand finds it optimal to recommend option h as the probability of continuing in case of fitting advice is at least β^+ which is the probability that makes him indifferent between recommending h and l.

3.1. Markov equilibrium

This section considers the most common refinement of Perfect Bayesian equilibrium – Markov equilibrium. One might hope to get additional, sharper results for this class of equilibria than for the wider class of all perfect Bayesian equilibria. Since Markov equilibrium is the predominant solution concept in the literature on reputation, such results will then allow a clearer comparison with prior work. An additional advantage of Markov strategies is that they are particularly simple and intuitive in the context of reputation since they depend only on the adviser's reputation α.

A mixed Markov strategy of DM is a function β that assigns to each reputation $\alpha \in [0,1]$ the probability with which DM continues the game. A mixed Markov strategy of A is a function s where $s(\alpha, b)$ is the probability of recommending option h when the bonus option is $b \in \{h, l\}$ and the reputation is α. A profile of Markov strategies that constitutes a perfect Bayesian equilibrium is called "Markov equilibrium".

As in other cheap talk games, a "babbling equilibrium" in which no meaningful advice is given always exists. Within the framework of this paper, the babbling equilibrium takes the following form: DM always ends the game and A always recommends his bonus option. This equilibrium is a Markov equilibrium in which the Markov strategies β and s are constant functions. Clearly, this is not the most interesting equilibrium. Mailath and Samuelson (2001) analyze a very similar model and their proposition 2.3 implies that strategies in a Markov equilibrium cannot be piecewise continuous unless A babbles.[7] They construct an equilibrium with an infinite number of discontinuities (proposition 2.2). However, the adviser's strategy in this equilibrium is equivalent to a babbling equilibrium for all but a countable number of beliefs α. This result holds true in the model of my paper as well and the underlying logic is briefly sketched in Appendix B. Hence, the Markov refinement essentially rules out meaningful communication.

As standard Markov equilibria do not allow for meaningful communication, I will in the remainder of this section extend the solution concept in a minimal way. The new solution concept, called quasi-Markov equilibrium, will still have the advantages of being simple and intuitive and will allow me to derive clear cut results for the existence of equilibria with meaningful communication as well as some properties of these equilibria. I define *quasi-Markov strategies* as follows: quasi-Markov strategies depend only on the adviser's reputation α and observations in the current period. In other words, a quasi-Markov strategy for A depends on the identity of the bonus option and the belief α which implies that quasi-Markov and Markov strategies are identical for A. For DM, however, the set of quasi-Markov strategies is a superset of the Markov strategies because quasi-Markov strategies depend not only on the reputation but also on whether the recommendation fitted DM's needs in the current period. A mixed strategy for DM is consequently denoted by two functions (β^+, β^-) where β^+ assigns to each $\alpha \in [0,1]$ the probability with which DM continues the game if the recommendation of the current period has fitted his needs. β^- assigns to each belief $\alpha \in [0,1]$ the probability with which DM continues if the recommendation did not fit DM's needs.[8] A profile of quasi-Markov strategies that constitutes a perfect Bayesian equilibrium is called "quasi-

[7] I call a function "piecewise continuous" if it is continuous at all but a *finite* number of points.

[8] It is convenient to let β^+ and β^- depend on the belief at the beginning of the period. This is also in line with the usual notation in Markov equilibrium. Note that this is simply a notational convention and having these functions depend on the updated belief would not yield different results because DM's strategy depends directly on whether the recommendation has fitted his needs or not.

Markov equilibrium". As before, I restrict myself to informative equilibria.

It is convenient to use value functions $V : [0,1] \to \mathbb{R}$ for A and $W : [0,1] \to \mathbb{R}$ for DM to describe quasi-Markov equilibria. More precisely, $V(\alpha)$ is – in a given quasi-Markov equilibrium – A's expected discounted payoff stream at the beginning of a period when A has reputation α. By "beginning" I refer to the point in time before A receives his signal, i.e. the expectation is also taken over A's signal. $W(\alpha)$ is DM's expected discounted payoff stream at the beginning of a period when A has reputation α.

The following lemma characterizes A's strategy in a quasi-Markov equilibrium. Unsurprisingly, A recommends option h if this option is the bonus option as in this case the interests of DM and A are aligned. If, however, the bonus option is option l, A has to trade-off dynamic and static incentives. He will then only recommend option h if the increase in future value of a fitting recommendation is sufficiently high.

Lemma 2. *Let V be A's value function in a quasi-Markov equilibrium. A's strategy in quasi-Markov equilibrium satisfies the following:*

1. *if option h is the bonus option, A recommends option h;*
2. *if option l is the bonus option, A recommends option h only if*

$$\beta^+(\alpha)V(\alpha^+) - \beta^-(\alpha)V(\alpha^-) \geq \frac{1}{\delta(2p-1)}; \tag{2}$$

and recommends option l only if the reverse inequality holds.

A is willing to mix if and only if (2) holds with equality.

DM's optimal strategy is relatively simple: he ends the game if his expected payoff (in $t+1$ and following periods) from continuing the game is lower than his outside option W_O. If the competent type recommends the option fitting DM's needs with probability q (given reputation α), then DM's value is

$$W(\alpha) = (\alpha q + (1-\alpha)/2)(1 + \delta W(\alpha^+)) + (\alpha(1-q) + (1-\alpha)/2)\delta W(\alpha^-). \tag{3}$$

As long as $W(\alpha) \geq W_O$, it is optimal to continue. DM is willing to use a mixed strategy if and only if $W(\alpha) = W_O$ and stops the game if $W(\alpha) < W_O$.

The starting point for introducing quasi-Markov equilibria was the observation that Markov equilibria with meaningful communication do not exist. The following result states a necessary and sufficient condition for the existence of quasi-Markov equilibria that satisfy a simple regularity condition and in which meaningful communication takes place for some reputation levels.

Definition 1. *A quasi-Markov equilibrium is* regular *if the strategies β^+, β^- and s are piecewise continuous and of bounded variation.*

Proposition 3. *If*

$$p \geq \frac{3}{4} + \frac{1-\delta}{2\delta}, \tag{4}$$

then there exists a regular quasi-Markov equilibrium in which $s(\alpha, l) > 0$ and $\beta^+(\alpha) > 0$ for all beliefs

$$\alpha > \underline{\alpha} = \frac{W_O(1-\delta) - 1/2}{p - 1/2}.$$

Conversely, if a regular quasi-Markov equilibrium exists in which $s(\alpha, l) > 0$ on some interval of reputations, then (4) holds. Furthermore, in such an equilibrium $W(\alpha) = W_O$ for $\alpha \in [\bar{\alpha}, 1]$ for some $\bar{\alpha} < 1$ and

$$\lim_{\alpha \to 1} V(\alpha) \in \left[\frac{4p-1}{4p-2}, \frac{4p-3}{(4p-2)(1-\delta)}\right]. \tag{5}$$

Condition (4) holds if A's signal is sufficiently informative and players are sufficiently patient. In particular, $p > 3/4$ and $\delta > 2/3$ are both implied by (4). This condition intuitively makes sense in so far as to incentivize A to give useful recommendations instead of cashing in on his bonus immediately, A must enjoy substantial payoffs in the future and he cannot be incentivized by future payoffs if he discounts those a lot. If the signal technology is bad, the welfare gains from advice are small. DM has to leave some of these welfare gains to the adviser in order to incentivize him; some rents go to the incompetent type and sometimes A will recommend his bonus option even if it is option l (theorem 1 implies that this has to be true in every equilibrium). If the signal technology is too bad, then the payoff left for DM is simply too small to prevent him from taking his outside option.

The second part of proposition 3 extends theorem 1. Since DM ends the game with positive probability for beliefs arbitrarily close to 1, DM's value has to be equal to his outside option value for reputations arbitrarily close to 1. The regularity of strategies then implies that $W(\alpha) = W_O$ on an interval of high reputation.

The last part of proposition 3 illustrates a basic commitment problem that A faces. Suppose A could commit to the strategy "always recommend option h". This would give DM the highest possible payoff and imply that DM does not stop the game if he believes with sufficiently high probability that A is competent. What is more surprising is that this commitment would also increase A's payoff. Note that the probability of recommending the bonus option would be 1/2 in each period implying that the expected payoff of A under commitment would be $1/(2(1-\delta))$. It is straightforward to verify that this payoff is higher than the upper bound of $\lim_{\alpha\to 1} V(\alpha)$ in (5).[9]

Corollary 1. *A's payoff in a regular quasi-Markov equilibrium is – for high values of reputation – lower than the commitment payoff of a competent type:* $\lim_{\alpha\to 1} V(\alpha) < 1/(2(1-\delta))$.

This section concludes with another example which describes a family of regular quasi-Markov equilibria. This family includes some examples where DM continues with probability 1 at some interior beliefs while A will still give informative advice at these beliefs due to dynamic incentives.

Example 2. *Roughly speaking there are three regions in terms of reputation: for reputations below α_1, A recommends the bonus option and DM fires the adviser for sure. If the reputation is between α_1 and α_2, then DM fires the adviser after receiving unfitting advice but continues with positive probability (usually less than 1) if the advice fits his needs. For reputations above α_2, DM continues for sure after receiving fitting advice but fires the adviser with (usually) positive probability after bad advice. This structure implies that A's value is relatively low for reputations between α_1 and α_2. Consequently, A is afraid of having his reputation drop into this range when his reputation is slightly above α_2 and this allows to sustain meaningful communication even if the probability of immediate firing is low. For certain parameter values, it is even possible to do without the threat of firing for α slightly above α_2, i.e. $\beta^-(\alpha) = 1$, without destroying A's truthtelling incentives completely.*

More precisely, take some α_1 and let

$$\alpha_2 = (2 - 2(1-\delta)W_O)\alpha_1 - 1 + 2(1-\delta)W_O \tag{6}$$

$$\alpha_3 = (2 - 2(1-\delta)W_O)\alpha_2 - 1 + 2(1-\delta)W_O \tag{7}$$

$$\tilde{\alpha} = 2(1-\delta)W_O\alpha_2 + 1 - 2(1-\delta)W_O \tag{8}$$

which will later ensure that $\alpha_2^- = \alpha_1$, $\alpha_3^- = \alpha_2$ and $\tilde{\alpha}^+ = \alpha_2$. Consider the following strategies and value functions:

[9]It should be noted that the bounds on the value function in (5) and consequently also corollary 1 hold only for regular quasi-Markov equilibria. The supplementary material illustrates that the results do not apply to all perfect Bayesian equilibria and derives bounds for A's expected discounted payoff stream that hold for all perfect Bayesian equilibria.

$$\beta^+(\alpha) = \begin{cases} 0 & \textit{if } \alpha < \alpha_1 \\ \frac{1}{\delta(2p-1/2)} & \textit{if } \alpha_1 \le \alpha < \tilde{\alpha} \\ \frac{1-\delta}{\delta(2p-3/2)} & \textit{if } \tilde{\alpha} \le \alpha < \alpha_2 \\ 1 & \textit{else} \end{cases} \qquad \beta^-(\alpha) = \begin{cases} 0 & \textit{if } \alpha < \alpha_2 \\ \frac{2p\delta-\delta/2-1}{\delta(1-\delta)(2p-1/2)} & \textit{if } \alpha_2 \le \alpha < \alpha_3 \\ \frac{2p\delta-\delta/2-1}{\delta(2p-3/2)} & \textit{else} \end{cases}$$

$$s(\alpha, l) = \begin{cases} 0 & \textit{if } \alpha < \alpha_1 \\ \frac{2(1-\delta)W_O-1}{\alpha(2p-1)} & \textit{else} \end{cases}$$

$$V(\alpha) = \begin{cases} 1 & \textit{if } \alpha < \alpha_1 \\ \frac{2p-1/2}{2p-1} & \textit{if } \alpha_1 \le \alpha < \alpha_2 \\ \frac{2p-3/2}{(1-\delta)(2p-1)} & \textit{else} \end{cases} \qquad W(\alpha) = \begin{cases} 1/2 + \delta W_O & \textit{if } \alpha < \alpha_1 \\ W_O & \textit{else} . \end{cases}$$

Proposition 4. *Assume $p \in [1/4 + 1/(2\delta), 1/4 + 1/(2\delta^2)]$ and $\alpha_1 \in [(2(1-\delta)W_O - 1)/(2p-1), 1)$. Then the strategies above constitute a quasi-Markov equilibrium. If $p = 1/4 + 1/(2\delta^2)$, then $\beta^-(\alpha) = \beta^+(\alpha) = 1$ on $[\alpha_2, \alpha_3)$.*

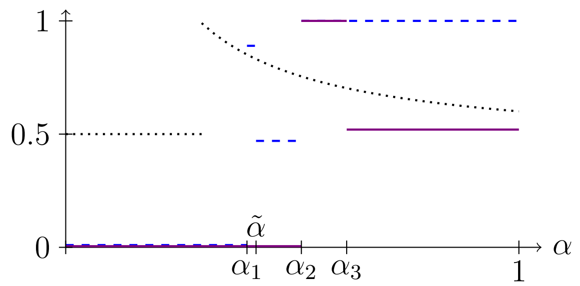

Figure 3: Example 2 for parameter values $\delta = 0.9$, $W_O = 6$, $p = 0.867$. (β^+ in dashed blue, β^- in solid violet, probability of fitting advice by a competent type in dotted black)

As shown in theorem 1, the property that no firing takes place on an interval of reputations can only occur for an interior interval. The way to interpret this example is as follows: For α below α_1, the adviser is fired due to perceived incompetence. For $\alpha \in [\alpha_1, \alpha_2)$, the adviser will be perceived as incompetent if he gives bad advice in this period. Because of this (and because the already skeptical DM is firing him with positive probability even after fitting advice), A's expected value is rather low for this belief range. For $\alpha \in [\alpha_2, \alpha_3)$, A knows that bad advice will reduce his reputation below α_2 where his value is low while fitting advice maintains a high value. Consequently, A is willing to give reasonably good advice although DM will not fire him even if this period's advice turns out to be bad. At high levels of reputation, $\alpha \ge \alpha_3$, having one's reputation drop below α_2 soon is sufficiently unlikely (and at least two periods away) and consequently

A would now recommend his bonus option if DM continued for sure. The threat of being fired due to incompetence soon is no longer sufficiently big to keep the moral hazard in check. As a result, DM has to fire A with positive probability after receiving advice that does not fit his needs.

3.2. Relaxing Assumptions

In this section, I want to discuss how some of the assumptions could be relaxed. I will keep the discussion short and informal but provide formal proofs generalizing theorems 1 and 2 in the supplementary material. Firstly, the assumption that A and DM share the same discount factor can be discarded without affecting any result qualitatively as long as the two discount factors are strictly less than 1. Secondly, one might wonder about a signal technology that is not constant over time; in other words, the posterior p of a competent type might depend on the specific period. Strategies should then naturally depend on the time period as well, implying that (quasi-) Markov strategies make less sense in this setting. The results on the (expected) length of the game, however, still hold true if one substitutes p by $\sup(\{p^t\})$ where p^t is the competent type's posterior in period t and the supremum is assumed to be strictly below 1. With this adjustment the proof of theorem 2 will go through and the results will hold. Thirdly, one can allow for more than two options at the expense of a somewhat more cluttered notation; see an earlier working paper version of this paper (Schottmüller, 2016).

Finally, I want to discuss the possibility of correlation between A's bonus option and the option fitting DM's needs. Firstly, consider within period correlation of the competent type's bonus option with DM's needs. As long as the correlation is imperfect, this will not change theorems 1 and 2 qualitatively because the only relevant part of the signal technology is A's posterior belief which was denoted by p. Correlation will now imply that the posterior p can depend on the identity of the bonus option. As long as the posterior is strictly between 1/2 and 1, the structure of the problem does not change and the proofs go through using the higher of the possible posteriors p. (Of course, the updating of DM's belief α will also be affected but this plays only a minor role.) Secondly, consider correlation over time. More precisely, the probability that A's bonus option coincides with the option fitting DM's needs may depend on whether it had also done so in the previous period.[10] Note that this form of correlation is irrelevant for A's decision problem because he knows his bonus option when making his recommendation (and the correlation does not improve the accuracy of his signal). The correlation will affect DM's beliefs and decisions. The proofs of theorems 1 and 2 are based on scenarios where A wants to recommend his bonus option although it is

[10] Here I still maintain the assumption that at the start of the period DM views both options as equally likely to fit his needs. Otherwise, the strategic effects discussed, for example, in Ottaviani and Sørensen (2006a,b) will play a role.

option l. Hence, as long as this scenario has positive probability – that is, so long as the correlation is not perfect – the same situation emerges and the results hold qualitatively.

4. Extensions

4.1. Transferable utility

So far, DM could not use monetary payments to incentivize the agent. One might conjecture that using monetary rewards for fitting advice could mitigate the moral hazard problem that leads to the firing of advisers with a high reputation. This section shows that theorems 1 and 2 continue to hold if monetary transfers from DM to A are allowed.

The model now allows DM to transfer a wage $w \geq 0$ to A at the end of each period t; i.e. at the same time when DM decides whether to continue or to stop. (The exact timing of the payments does not matter for the results.) A's payoff in a given period is $1 + w$ if he recommends his bonus option and w else where w is the wage payment he receives. DM's payoff in a given period is $1-w$ if the recommendation fits his needs and $-w$ else. In the spirit of relational contracting, I will not allow the players to commit to an enforceable payment schedule but will consider self enforcing transfers only.[11]

Theorems 1 and 2 remain true in this setting. To see this, it is useful to recall the main idea behind the proof of theorem 1. The proof was by contradiction. Hence, suppose DM never fired the adviser if his belief is above some threshold $\bar{\alpha} < 1$. For every $T \in \mathbb{N}$ one can then find a belief α_T sufficiently close to 1 such that an adviser with reputation α_T will maintain a reputation above $\bar{\alpha}$ even if A's recommendations do not fit DM's needs for T consecutive periods. An adviser with reputation α_T (or higher) can therefore guarantee himself a payoff of $\sum_{t=0}^{T} \delta^t$ by recommending his bonus option T times. Hence, his equilibrium payoff must be at least $\sum_{t=0}^{T} \delta^t$. For sufficiently high T one can show that $\sum_{t=0}^{T} \delta^t$ is higher than the welfare that the advice relationship generates. Consequently, DM would be obliged to have an expected payoff below the payoff he would obtain by immediately stopping the game which is impossible in equilibrium.

Note that the lower bound on A's payoff in the previous argument remains valid because the wage w cannot be negative. Put differently, wage payments can only increase A's payoff to even more than $\sum_{t=0}^{T} \delta^t$. Welfare is also not affected by the introduction of a lump sum transfer. Consequently, the argument still holds and the proof of theorem 1 goes through without substantial changes. The proof of theorem 2 is based on a similar lower bound for A's payoff and therefore also continues to hold.

[11] With commitment, the usual way of solving moral hazard problems – selling the enterprise to the agent – will work. Therefore, in the context of this paper, this would entail that DM offers A to commit to paying him 1 unit for every fitting recommendation if A pays $p/(1-\delta) - \varepsilon$ to DM ex ante. The competent type will accept the offer and the incompetent will not.

There are, however, two caveats that I briefly mention here and illustrate more extensively in the supplementary material. Firstly, the results do not hold if A can also make payments to DM. In this case, there are equilibria in which both types of the adviser will compensate DM for bad recommendations and DM will continue forever, regardless of which type he is facing. Such equilibria exist because, from a total welfare point of view, it is efficient for DM to interact even with an incompetent type: the incompetent type's bonus payments are more valuable than DM's outside option by (1). It is then unsurprising that the efficient outcome can be obtained with payments from A to DM if players are sufficiently patient.

Secondly, in order to to derive the results it was assumed that A's gain from recommending his bonus option was as high as DM's payoff from receiving a fitting recommendation – both were assumed to be 1. Equilibria in which A recommends option h even when his reputation is high can be sustained if A's payoff from recommending the bonus option is sufficiently small. Intuitively, A could be incentivized to always recommend option h if he received an incentive payment of $\kappa/(2p-1)$ for fitting recommendations where κ denotes the payoff that A can get from recommending his bonus option. For κ sufficiently small, DM would be willing to make this incentive payment to obtain the best possible advice from a competent type. Consequently, the results above only hold if κ is not too small.

4.2. Competition among advisers

One can easily imagine situations in which a decision maker has to decide which of several advisers to consult. In this vein, DM's outside option in the model could be interpreted as the value he assigns to advice from the next best adviser. This section briefly discusses two settings that reflect this idea and shows how the results of the paper can be reinterpreted in these alternative model settings.[12]

Assume that there is a finite number of potential advisers – for simplicity, say two. DM's initial belief α_0^i that adviser i is competent might differ between these two advisers. Both advisers have the same preferences and signal technology as in the main model and their types are assumed to be uncorrelated. Furthermore, DM can ask only one adviser each period. As a first setting, assume that DM still has an exogenous outside option satisfying (1). At the end of a period DM has then three actions: *stop* the game and take the outside option, *stay* with the current adviser or *switch* to the other adviser.

In this setting, there are direct analogues of theorems 1 and 2. The reason is that both results were proven by showing that A would prefer to recommend his bonus option with probability 1 if the conditions of the theorems were not satisfied (and in this case DM prefers stopping to staying with this adviser). Note that A is even

[12] A third possible setting is discussed in the supplementary material.

more tempted to recommend his bonus option in the current setting because there is a possibility that instead of stopping the game, DM may switch and then switch back later. Consequently, the proofs still apply. Theorem 1 implies that for a fixed α^{-i}, there is no $\bar{\alpha}$ such that DM always stays with adviser i if $\alpha^i > \bar{\alpha}$. Since this holds for any given and fixed α^{-i}, the statement is also true without conditioning on α^{-i}: there is no $\bar{\alpha}$ such that DM chooses adviser i whenever $\alpha^i > \bar{\alpha}$. Similarly, theorem 2 still holds but now states that DM will end the game *or switch* (possibly to an adviser with a worse reputation) within the given number of periods. Let me also point out the dual nature of the inefficiency established in this paper: First, the decision maker does not get the best possible advice, and second, he will eventually switch from his (originally) most preferred adviser to less preferred ones or he will stop asking for advice altogether.

As a second setting, consider the same setting as above, but eliminate the exogenous outside option and allow for n instead of 2 advisers. In other words, DM has the choice between *staying* with the current adviser or *switching* which means that he will randomly be matched with another adviser from the pool of n advisers. DM's outside option is therefore completely endogenized. Note that this game still has a babbling equilibrium in which all advisers always recommend their bonus option and DM's strategy is independent of whether the recommendations have fitted his needs or not, for example, DM always stays with the first adviser. In a literal sense, this babbling equilibrium violates theorem 1 since DM always stays with the first adviser. However, the following analogue to theorem 1 is directly implied by the proof of theorem 1: if in a given equilibrium there exists an $\bar{\alpha}$ such that DM stays with adviser i with probability 1 for all $\alpha^i > \bar{\alpha}$, then babbling emerges after all histories that lead to a sufficiently high reputation for adviser i.[13] A similar analogue to theorem 2 would state that adviser i will always recommend his bonus option if the expected length of the advice relationship is above $\bar{T}$ after some history.

5. Related literature

Related to this paper is the literature on cheap talk started by Crawford and Sobel (1982) and surveyed in Krishna and Morgan (2008).[14] The exact structure of the payoffs is, however, somewhat different from the traditional cheap talk setup where

[13]Following the argument in the proof of theorem 1, this result could be strengthened with regard to the case where DM stays with adviser i with probability of at least $1-(\eta-\eta\delta)/(2-2\eta\delta)$ whenever $\alpha^i > \bar{\alpha}$. As in the proof of theorem 1, this would imply that adviser i will always recommend his bonus option for sufficiently high α.

[14]Advice is in my setup directly relevant for the decision maker's payoff but the main reason is that – due to his ignorance – the decision maker has no real choice but to follow the adviser's advice (as long as he has not quit the advice relationship). Put differently, the model of this paper is similar to a model in which the advice is real cheap talk and the decision maker has a pseudo decision on whether to follow the advice or not; see the supplementary material to this paper for a more detailed discussion.

the adviser has a bias in a commonly known direction whereas in my setup the bias is for one random option that carries a bonus for the adviser. Within the cheap talk literature, models of repeated cheap talk in which the state of the world changes in each period are closest to my paper. Park (2005) analyzes a situation where a consumer has a problem in each period and relies upon advice to find out which of the several repair shops specializes in fixing the problem at hand. In contrast to the current paper, the adviser knows the state of the world perfectly; consequently, reputation for competence does not play a role. Hence, an expert cannot be "too good" which is the main focus of my paper. This is also the main difference to earlier papers (Sobel, 1985; Benabou and Laroque, 1992) where the adviser's type refers to his honesty and not the quality of his information. Being more certain of facing an honest type is good for consumers, implying that they will certainly not fire the adviser at these favorable beliefs. In my model, being more likely to face an informed type can be bad because it can aggravate the moral hazard problem.

An effect closely related to this paper is known from the literature on reputation with repeated interaction between a sequence of short lived principals and a long lived agent where both moral hazard and adverse selection are present, see Mailath and Samuelson (2001) and Jullien and Park (2014). (In the following, I refer to Jullien and Park's section 6.1 which is closest to my setup since the agent is uncertain). In these papers, incentivizing agents with a good reputation becomes so difficult that equilibria with high effort (where high effort corresponds to truthfully revealing one's signal in Jullien and Park's model) do not exist if monitoring is noisy.[15] The reason is – just like in my paper – that the belief updating is slow when the belief of facing a competent type is close to 1. However, a high effort equilibrium may exist if types are impermanent; that is, in each period the type of the agent changes – unobserved by the principal – with a small probability. A similar logic also appears in Cripps et al. (2004) and Wiseman (2008). My paper differs in several ways: Firstly, I do not restrict myself to Markov equilibria because these do not allow for information transmission, see section 3.1. Using, for example, quasi-Markov equilibria allows for partially informative equilibria that still display the difficulty of incentivizing advisers with high reputation: advisers with better reputation give worse advice and are fired with some probability. In contrast, Mailath and Samuelson (2001) and Jullien and Park (2014) focus on "honest Markov equilibria" in which good agents exert high effort (Mailath and Samuelson, 2001) or truthfully reveal their information (Jullien and Park, 2014) regardless of their reputation and a higher reputation is therefore always good news for the principal.

[15] One exception is proposition 2 in Mailath and Samuelson (2001) where a mixed equilibrium is constructed in which high effort is exerted for some reputation levels. However, the set of these reputation levels is countable and strategies necessarily have an infinite number of discontinuities in any such equilibrium; see section 3.1 for a more detailed discussion.

Furthermore, types are permanent in my model (which would lead to the non-existence of honest Markov equilibria in the papers mentioned above) and both the adviser and the decision maker are long lived and strategic. Having a long lived decision maker allows for a focus on stopping/firing which is less relevant in the aforementioned papers with short lived principals.

More broadly, this paper is part of the literature asking whether career concerns and reputation can prevent opportunism, see Fama (1980) and Holmström (1982) for seminal contributions. Closest is Aghion and Jackson (2016) in which (political) “leaders” have to be incentivized to take risky decisions (instead of remaining inactive) by the threat to vote them out of office. In equilibrium, even arbitrarily competent leaders are terminated with some probability whenever they do not take a risky decision. However, setup and applications differ significantly since “leaders” do not receive bonuses and are unaware of their own type.

Several papers in the literature on experts and advice, e.g. Brandenburger and Polak (1996); Ottaviani and Sørensen (2006a,b), analyze how an expert who wants to maximize his reputation for being competent will misrepresent his information. The main result is that the adviser will then misrepresent his signal towards the prior. This effect is not present in the current paper because the decision maker has a uniform prior which makes it impossible to misrepresent towards the prior. The main difference, however, is that the expert in my model wants to maximize his expected bonus stream – and not his reputation per se – which leads to the moral hazard problem that drives the results of my paper.

6. Conclusion

This paper has analyzed the question why advisers are fired. Two reasons for their dismissal are identified in a repeated game model. Firstly, incompetence; that is, advisers who are believed to be of low quality are fired. Secondly, (justified!) mistrust. Advisers who are believed to be competent are not afraid of being fired due to incompetence. In equilibrium, these advisers will therefore push their own agenda, i.e. recommend actions that contribute to their own benefit more than to the decision maker’s benefit, unless they face the threat of dismissal. Consequently, the decision maker has to fire them with positive probability in equilibrium (when receiving bad advice). The interplay of these two effects can imply that the decision maker receives the most informative advice from an adviser whose level of competence is unclear. Such an adviser tries to give good advice because he will be perceived as being of such a low quality that he is fired due to incompetence in the event that his advice turns out to be bad. The firing of competent advisers is inevitable in equilibrium but inefficient. Independently of competence and beliefs, the expected length of the advice relationship is therefore limited even though advice by a competent adviser is efficient. Thus, the existence of private benefits for the adviser, such as bonus payments, does not only lead to bad advice, but also implies that decision makers eventually drop the best advisers and end up receiving inferior (or no) advice.

There is a caveat to the results above: if the decision maker has additional instruments with which he can punish advisers who give bad advice, he may be able to ensure good advice without the need to fire. However, in many situations not seeking advice in the future is the only instrument available to the advice seeker. If a customer is unhappy with the advice of his bank or a store's sales personnel, or a patient suspects that his general practitioner is being influenced by lobbying efforts of the pharmaceutical industry, or a politician mistrusts his advisers, the only possible remedy will often be to stop listening to the existing advisers (and maybe switch to new ones). This is even truer in case of everyday informal advice given to people by colleagues or friends.

The model of this paper helps to identify the effects mentioned above and paves the way for further research. For example, the literature on sell-side analysis in financial advice – see, for example, Fang and Yasuda (2009); Jackson (2005) – is concerned with the extent to which reputation can alleviate opportunistic behavior by analysts.[16] While this literature establishes empirically that more reputable analysts provide better predictions on average, this result is based on the assumption of either a binary or a linear functional form for this relationship. However, the model of this paper suggests the possibility of a non-monotonic (inversely U-shaped) relationship.[17] On the theoretical side, one might explore instruments that a decision maker could use to discipline advisers, such as own (costly) acquisition of noisy information, or consulting multiple advisers concurrently. Another interesting question worth raising is the effect of learning. Here, an adviser might receive more precise signals concerning the decision maker's needs as he interacts with the decision maker repeatedly. These and other possibilities are beyond the scope of the current paper and are left for future research.

[16] "Sell-side analysis" refers to situations in which employees of a broker provide analysis and stock recommendations to potential customers for free in the hope of generating an order that yields a commission.

[17] The possibility of non-monotonicities seems to have escaped the attention of the authors, e.g. Fang and Yasuda (2009, p. 3736) write "Because analysts with a better reputation have greater long-term benefits to lose, theory predicts that they are more likely to refrain from opportunism." My paper has demonstrated that this argument, though plausible at first sight, may in fact not be true in equilibrium.

Appendix A: Proofs for perfect Bayesian equilibria

Proof of lemma 1: Recommending the bonus option yields an expected payoff of $1 + v^+/2 + v^-/2$ where v^+ (v^-) denotes the expected discounted future payoff stream after a fitting (non-fitting) recommendation. Recommending the non-bonus option has expected payoff $v^+/2 + v^-/2 < 1 + v^+/2 + v^-/2$. □

Proof of proposition 1: DM's expected payoff from continuing is bounded from above by $\delta\alpha p/(1-\delta) + \delta(1-\alpha)/(2(1-\delta))$. For α sufficiently low (but strictly higher than 0), this upper bound is less than δW_O as $W_O > 1/(2(1-\delta))$ by (1). □

Proof of theorem 1: The proof is by contradiction. Suppose – contrary to the theorem – that there exists an $\bar{\alpha} < 1$ such that DM continues with probability of at least $1-\varepsilon$ whenever his belief is above $\bar{\alpha}$. As belief updating becomes arbitrarily slow when α is close to 1, there exists an α_T for every $T \in \mathbb{N}$ such that an adviser of reputation α_T (or higher) will still have a reputation above $\bar{\alpha}$ after T periods, i.e. even if he recommends for T consecutive times a non-fitting option. An adviser with reputation of (at least) α_T can guarantee himself a continuation payoff of

$$\hat{V} = \sum_{t=0}^{T} (\delta(1-\varepsilon))^t = \frac{1-\delta^{T+1}}{1-\delta(1-\varepsilon)}$$

by recommending his bonus option for the next T periods. Note that this is true regardless of A's type. Consequently, even if we take expectations over A's type his equilibrium continuation profit is at least $\hat{V}$ if his reputation is above α_T.

Next, an upper bound for the sum of expected continuation payoffs is derived. The sum of expected payoffs in a given period (after an arbitrary history and when taking expectations at the start of the period before A's signal realizes) has to be less than $1/2 * 1 + 1/2 * 2$ as with probability 1/2 the bonus option coincides with the option fitting DM's needs – allowing for a maximal possible sum of payoffs of 2 – and with probability 1/2 the two do not coincide which means that only one of the two players can have a payoff of 1. Consequently, the sum of expected continuation payoffs after any given history has to be less than $(3/2)/(1-\delta)$.

DM can ensure himself an expected payoff of at least $1/2 + \delta W_O$ after any given history by stopping the game at the end of the current period. (1/2 is the payoff DM expects to receive from an incompetent type in a given period, and by informativeness of the equilibrium, 1/2 is also the minimum expected payoff from a competent type.) The upper bound on continuation welfare and the lower bound on DM's continuation payoff yield an upper bound on A's expected continuation payoff equal to

$$\bar{V} = \frac{3/2}{1-\delta} - 1/2 - \delta W_O = \frac{1-\eta\delta}{1-\delta} \tag{9}$$

where $\eta = (1-\delta)W_O - 1/2$ and $\eta > 0$ by (1).[18]

For T sufficiently high, however, the lower bound on A's equilibrium continuation payoffs $\hat{V}$ is above the upper bound for A's equilibrium continuation payoffs $\bar{V}$:

$$\begin{aligned} \frac{1-\delta^{T+1}}{1-\delta(1-\varepsilon)} &> \frac{1-\eta\delta}{1-\delta} \\ \Leftrightarrow 1-\delta-(1-\delta)\delta^{T+1} &> 1-\eta\delta-\delta(1-\varepsilon)+\eta\delta^2(1-\varepsilon) \\ \Leftrightarrow (1-\varepsilon)(1-\eta\delta) &> 1-\eta+(1-\delta)\delta^T \\ \Leftrightarrow \frac{(\eta-\delta^T)(1-\delta)}{1-\eta\delta} &> \varepsilon \end{aligned}$$

which is true by assumption for sufficiently high T and for $\varepsilon = \eta(1-\delta)/(2-2\eta\delta) > 0$. $\hat{V} > \bar{V}$ contradicts that the supposed equilibrium exists and therefore establishes the result. □

Proof of theorem 2: This proof will use the upper bound $\bar{V}$ on A's continuation payoff derived in the proof of theorem 1, see equation (9).

First, it is shown by contradiction that – after any given history – DM stops the game with at least probability ε' along some path of length T'. Suppose this was not the case. A can then guarantee himself an expected continuation payoff of

$$\underline{V} = \frac{1-\delta^{T'+1}}{1-\delta(1-\varepsilon')}$$

by recommending the bonus option for the next T' periods. (This – very conservative – lower bound assumes a worst case where DM stops the game with probability ε' in each of the following T' periods and stops the game with probability 1 in period $T'+1$. This lower bound on A's continuation payoff holds regardless of A's type and holds therefore also if expectation is taken over A's type.) The contradiction emerges as $\underline{V} > \bar{V}$:

$$\begin{aligned} \frac{1-\delta^{T'+1}}{1-\delta(1-\varepsilon')} &> \frac{1-\eta\delta}{1-\delta} \\ \Leftrightarrow 1-\delta-(1-\delta)\delta^{T'+1} &> 1-\delta\eta-\delta(1-\varepsilon')+\delta^2(1-\varepsilon')\eta \\ \Leftrightarrow (1-\delta\eta)(1-\varepsilon')-(1-\eta) &> (1-\delta)\delta^{T'} \\ \Leftrightarrow T' &> \frac{log\left(\frac{(1-\delta\eta)(1-\varepsilon')-1+\eta}{1-\delta}\right)}{log(\delta)} \end{aligned}$$

which holds true given the values of T' and ε' stated in the theorem. Hence, the game ends with at least ε' probability along some path of length T' in equilibrium (starting from any given history). Note that every path of length T', i.e. a sequence of T' fitting

[18] Note that for this upper bound on A's expected continuation payoff as well as for the lower bound on DM's continuation payoff above the expectation is taken also over A's type.

or not fitting recommendations starting after a given history h, has at least probability $\gamma = (1-p)^{T'}$ (conditional on h). This implies that the probability that DM continues for mT' periods is at most $(1-\gamma\varepsilon')^m$. Now take an arbitrary $\varepsilon > 0$ and let $m \in \mathbb{N}$ be sufficiently high such that $(1-\gamma\varepsilon')^m \leq \varepsilon$, e.g. $m = \lceil log(\varepsilon)/log(1-\gamma\varepsilon')\rceil$. Then, the probability that the game continues for T_ε or more periods is less than ε. This establishes the first result.

The second result follows from the first result: I just established that the probability that the game lasts longer than T_ε periods is *at most* ε. As I want to derive an upper bound on the expected length, I can assume that the probability that the game lasts longer than T_ε periods is *exactly* ε. As it simplifies the derivation and since I am only interested in an upper bound, I will actually assume that the probability that the game lasts longer than

$$\tilde{T}_\varepsilon = \frac{\log(\varepsilon)}{\log\left(1-(1-p)^{T'}\varepsilon'\right)}T' + T'$$

equals ε for $\tilde{T}_\varepsilon > T'$, which again will increase the expectation as $\tilde{T}_\varepsilon \geq T_\varepsilon$. That is, I assume that the game lasts at least T' periods (which again increases the expectation). Rearranging yields that the probability that the game's length is $\hat{T} > T'$ or less is $1 - e^{(\hat{T}-T')/B}$ where $B = T'/\log\left(1-(1-p)^{T'}\varepsilon'\right)$. Note that $B < 0$. The corresponding density is $-e^{(\hat{T}-T')/B}/B$. This allows to compute an upper bound on the expected length of the game as

$$\begin{aligned} T' + \int_{T'}^{\infty} -\frac{\hat{T}e^{(\hat{T}-T')/B}}{B}\,\mathrm{d}\hat{T} &= T' + \left[-\hat{T}e^{(\hat{T}-T')/B} + Be^{(\hat{T}-T')/B}\right]_{T'}^{\infty} \\ &= 2T' - B = T'\left(2 - \frac{1}{\log(1-(1-p)^{T'}\varepsilon')}\right). \end{aligned}$$

□

Proof of proposition 2: First, note that off-path behavior is optimal for both players: Given that DM ends the game it is optimal for A to recommend his bonus option and, given this, it is optimal to end the game because of (1). (This is basically the babbling equilibrium from the cheap talk literature.) The same argument establishes that players play optimal for $\alpha < \underline{\alpha}$. Note furthermore that this off-path behavior ensures that DM wants to end the game after a non-fitting recommendation if $\alpha \leq \bar{\alpha}$ as A will only recommend his bonus option in all future periods in case DM continues.

For $\alpha > \bar{\alpha}$, the expected discounted payoff stream of DM equals W_O as $s(\alpha, l)$ is such that

$$W_O = \alpha\left(\frac{1}{2}p + \frac{1}{2}\left(s(\alpha,l)p + (1-s(\alpha,l))(1-p)\right) + \delta W_O\right) + (1-\alpha)\left(\frac{1}{2} + \delta W_O\right).$$

Hence, DM is indifferent between ending the game and continuing after a fitting recommendation and therefore his strategy is a best response (given $\alpha > \bar{\alpha}$). Note that by the parameter condition $p > (1 + 2/\delta)/4$, $\beta^+(\alpha) < 1$, i.e. DM's strategy is indeed a feasible mix. Next I will show that A's expected discounted payoff stream for $\alpha > \bar{\alpha}$ equals $\tilde{V} = (2p - 1/2)/(2p - 1)$. Note that β^+ is such that A is indifferent between recommending options l and h if his bonus option is l (given the continuation value $\tilde{V}$) as

$$1 + \delta(1 - p)\beta^+\tilde{V} = \delta p\beta^+\tilde{V} \tag{10}$$

by the definition of β^+ and $\tilde{V}$. Hence, A's expected discounted payoff stream can be calculated as if A recommended always his bonus option, i.e.

$$\tilde{V} = 1 + \frac{1}{2}\delta p\beta^+\tilde{V} + \frac{1}{2}\delta(1 - p)\beta^+\tilde{V},$$

which holds by the definition of $\tilde{V}$ and β^+. As A is indifferent between recommending l and h when his bonus option is l, his strategy is a best response for $\alpha > \bar{\alpha}$.

For $\alpha \in [\underline{\alpha}, \bar{\alpha}]$, note that A's expected discounted payoff stream is at least $\tilde{V}$: $\tilde{V}$ is A's expected discounted payoff stream for $\alpha > \bar{\alpha}$ and for these beliefs this value $\tilde{V}$ can be achieved by always recommending option h (recall that β^+ is such that A is indifferent). For $\alpha \in [\underline{\alpha}, \bar{\alpha}]$, always recommending option h has to achieve a weakly higher expected payoff stream than $\tilde{V}$ as DM continues with probability 1 instead of $\beta^+ \leq 1$ in case of a fitting recommendation and $\alpha^+ \leq \bar{\alpha}$ while DM continues with probability 0 after a non-fitting recommendation both for α above and for α below $\bar{\alpha}$. Denoting A's expected discounted payoff stream by $V(\alpha)$, it follows that recommending h in case the bonus option is l is optimal for A as $\delta pV(\alpha^+) \geq 1 + \delta(1-p)V(\alpha^+)$ by (10) and $V(\alpha^+) \geq \tilde{V}$. Hence, A's strategy is indeed a best response. For DM, continuing is a best response as A will give best possible advice next period and therefore DM's payoff from continuing is at least $\delta\left(\alpha p + (1 - \alpha)/2 + \delta W_O\right)$ which is higher than δW_O by $\alpha \geq \underline{\alpha} > ((1-\delta)W_O - 1/2)/(p - 1/2)$. Consequently, DM's strategy is indeed a best response. □

Appendix B: Markov and quasi-Markov equilibria

Markov equilibria

It is convenient to use value functions $V : [0, 1] \to \mathbb{R}$ for A and $W : [0, 1] \to \mathbb{R}$ for DM to describe Markov equilibria. More precisely, $V(\alpha)$ is – in a given Markov equilibrium – A's expected discounted payoff stream at the beginning of a period when A has reputation α. With "beginning" I refer to the point in time before A receives his signal, i.e. the expectation is also taken over A's signal. $W(\alpha)$ is DM's expected discounted payoff stream at the beginning of a period when A has reputation α.

I will briefly sketch the argument why no Markov equilibrium with piecewise continuous strategies apart from the babbling equilibrium exists, see Mailath and Samuelson (2001) for a more detailed discussion. Note that A will recommend option h when his bonus option is l only if

$$p\beta(\alpha^+)\delta V(\alpha^+) + (1-p)\beta(\alpha^-)\delta V(\alpha^-) \geq 1 + (1-p)\beta(\alpha^+)\delta V(\alpha^+) + p\beta(\alpha^-)\delta V(\alpha^-).$$

That is, only if $\beta(\alpha^+)V(\alpha^+) - \beta(\alpha^-)V(\alpha^-) \geq 1/(\delta(2p-1))$. Consider now $\beta(\alpha^+)V(\alpha^+) - \beta(\alpha^-)V(\alpha^-)$ as $\alpha \to 1$. With piecewise continuous strategies V and β will be continuous for all α between a certain threshold and 1. From Bayes' rule, it then follows that $\alpha^+ - \alpha^- \to 0$ as $\alpha \to 1$. Together with the continuity of V and β, this implies that $\beta(\alpha^+)V(\alpha^+) - \beta(\alpha^-)V(\alpha^-)$ converge to 0 as $\alpha \to 1$. Consequently, A will recommend his bonus option regardless of his signal for beliefs above some threshold $\hat{\alpha} < 1$. Now suppose that – for some lower belief – A recommended option h with positive probability if option l is his bonus option. Looking at the highest such belief α, it is evident that DM should stop the game after a fitting recommendation as (i) $\alpha^+ > \alpha$ and (ii) A always recommends his bonus option for all beliefs above α. But in this case it is certainly not optimal for A to recommend option h if option l is his bonus option and his reputation is α which contradicts the fact that he does so with positive probability. Hence, no such α exists and babbling is the only Markov equilibrium in piecewise continuous strategies.

Quasi-Markov equilibria

Concentrating on informative equilibria allows to state the following technical result which establishes that A's expected future payoff stream is more valuable after a fitting than a non-fitting recommendation.

Lemma 3. *Let V be A's value function in a quasi-Markov equilibrium. Then, $\beta^+(\alpha)V(\alpha^+) \geq \beta^-(\alpha)V(\alpha^-)$.*

Proof of lemma 3: Suppose to the contrary that $\beta^+(\alpha)V(\alpha^+) < \beta^-(\alpha)V(\alpha^-)$. Then, A has an incentive to recommend options that do not fit DM's needs as this will give him the higher continuation value. Consequently, A will recommend either his bonus option or option l. Hence, a competent adviser will give (weakly) worse advice than an uninformed adviser which implies $\alpha^+ \leq \alpha^-$. In case of $\alpha^+ = \alpha^-$ Bayesian updating implies $\alpha^+ = \alpha^- = \alpha$, i.e. updating stops at belief α which implies that competent and incompetent both recommend their bonus option at belief α (regardless of whether the bonus option is l or h). By assumption (1), DM's best response to this behavior is to stop the game. That is $\beta^-(\alpha) = \beta^+(\alpha) = 0$ in a quasi-Markov equilibrium which implies that $\beta^+(\alpha)V(\alpha^+) < \beta^-(\alpha)V(\alpha^-)$ cannot hold. Finally, in case $\alpha^+ < \alpha^-$ the equilibrium is not informative. $\square$

Proof of lemma 2: The first item is shown first. A's expected utility in a given period can be written as

$$\delta q\beta^+(\alpha)V(\alpha^+) + \delta(1-q)\beta^-(\alpha)V(\alpha^-) + \mathbb{1}_{bonus}$$

where q is the probability that the recommendation satisfies DM's needs (which depends on the specific recommendation) and $\mathbb{1}_{bonus}$ is the indicator function for the bonus option, i.e. it is 1 if A recommends the bonus option and 0 otherwise. Lemma 3 (just above) states that $\beta^+(\alpha)V(\alpha^+) \geq \beta^-(\alpha)V(\alpha^-)$ in an informative equilibrium which implies that A will always recommend option h if option h is the bonus option: This recommendation maximizes the chance of improving his reputation and also pays him a bonus.

For the second item, let option l be the bonus option. Recommending option h will then give an expected payoff of $p\,\delta\beta^+(\alpha)V(\alpha^+)+\delta(1-p)\beta^-(\alpha)V(\alpha^-)$ and recommending option l will yield an expected payoff of $(1-p)\,\delta\beta^+(\alpha)V(\alpha^+)+\delta p\beta^-(\alpha)V(\alpha^-)+1$. The former is higher than the latter if and only if (2) holds. $\square$

Proof of proposition 3:

Sufficiency of (4): The proof is by construction. Let (4) be satisfied and define

$$\begin{aligned} \underline{\alpha} &= \frac{W_O(1-\delta)-1/2}{p-1/2}, \\ \check{V} &= \frac{4p-1}{4p-2}. \end{aligned}$$

Note that $\underline{\alpha} < 1$ by (1). Furthermore, $\check{V} > 1$ as $p > 3/4$ by (4).

I will show that the following value functions and strategies constitute a Markov equilibrium:

$$
\begin{aligned}
V(\alpha) &= \begin{cases} 1 & \text{if } \alpha \leq \underline{\alpha} \\ \check{V} & \text{if } \alpha > \underline{\alpha} \end{cases} \\
W(\alpha) &= \begin{cases} 1/2 + \delta W_O & \text{if } \alpha \leq \underline{\alpha} \\ W_O & \text{if } \alpha > \underline{\alpha} \end{cases} \\
s(\alpha, h) &= 1 \\
s(\alpha, l) &= \begin{cases} 0 & \text{if } \alpha \leq \underline{\alpha} \\ \frac{(1-\delta)W_O - 1/2}{\alpha(p-1/2)} & \text{if } \alpha > \underline{\alpha} \end{cases} \\
\beta^-(\alpha) &= 0 \\
\beta^+(\alpha) &= \begin{cases} 0 & \text{if } \alpha \leq \underline{\alpha} \\ \frac{1}{\check{V}\delta(2p-1)} & \text{if } \alpha > \underline{\alpha} \end{cases}
\end{aligned}
$$

where $s(\alpha, b)$ gives the probability with which A recommends option h if option b is his bonus option. Note that $0 \leq 1/(\check{V}\delta(2p-1)) \leq 1$ by (4).

The above is a quasi-Markov equilibrium if and only if the strategies are mutual best responses and the value functions are consistent with the strategies; i.e. if and only if the Bellman equations

$$
\begin{aligned}
V(\alpha) &= \max_{s_h, s_l} \Big\{ \frac{1}{2} \left[s_h + \delta q^+(s_h)\beta^+(\alpha)V(\alpha^+) + \delta q^-(s_h)\beta^-(\alpha)V(\alpha^-) \right] \\
&\quad + \frac{1}{2} \left[(1 - s_l) + \delta q^+(s_l)\beta^+(\alpha)V(\alpha^+) + \delta q^-(s_l)\beta^-(\alpha)V(\alpha^-) \right] \Big\} \\
W(\alpha) &= \max_{\beta^+, \beta^-} \big\{ q^+(\alpha) + \delta q^+(\alpha)\beta^+ W(\alpha^+) + \delta q^-(\alpha)\beta^- W(\alpha^-) \\
&\quad + \delta(1 - q^+(\alpha)\beta^+ - q^-(\alpha)\beta^-) W_O \big\}
\end{aligned}
$$

hold and the strategies $s(\alpha, l)$, $s(\alpha, h)$, $\beta^-(\alpha)$, $\beta^+(\alpha)$ are the respective maximizing arguments (where $q^+(s_i) = s_i p + (1 - s_i)(1 - p)$ is the probability of giving a fitting recommendation and $q^-(s_i) = 1 - q^+(s_i)$ is the counter-probability; similarly and with a slight abuse of notation, $q^+(\alpha) = (1 - \alpha)/2 + \alpha[s(\alpha, h)p + (1 - s(\alpha, h))(1 - p)]/2 + \alpha[s(\alpha, l)p + (1 - s(\alpha, l))(1 - p)]/2$ is the expected probability of getting a fitting recommendation and $q^-(\alpha) = 1 - q^+(\alpha)$). Consistency of the value functions with the strategies is straightforward to verify, and I will therefore only briefly explain why the strategies are indeed the maximizing arguments.

Recall that A wants to recommend option h in case the bonus option is option l if and only if $\beta^+(\alpha)V(\alpha^+) - \beta^-(\alpha)V(\alpha^-) \geq 1/(\delta(2p-1))$, see lemma 2. DM's strategy is chosen such that A is indifferent for $\alpha > \underline{\alpha}$. For $\alpha \leq \underline{\alpha}$, recommending the bonus option is clearly optimal as $\beta^+(\alpha) = \beta^-(\alpha) = 0$.

DM acts optimally when $\alpha \leq \underline{\alpha}$ as in this case $q^+(\alpha) = 1/2 = q^-(\alpha)$ and $\alpha^+ = \alpha^- = \alpha$ and therefore stopping the game gives W_O (in the next period) while continuing gives $1/2 + \delta W_O$ which is less than W_O by (1). For $\alpha > \underline{\alpha}$, $s(\alpha, l)$ is chosen such that DM is indifferent between continuing and stopping after a fitting recommendation (note that $q^+(\alpha) = (1-\delta)W_O$). After a non-fitting recommendation, DM is either indifferent (if $\alpha^- > \underline{\alpha}$) or strictly prefers stopping (if $\alpha^- \leq \underline{\alpha}$).

Necessity of (4): This proof is split up into two lemmas:

Lemma 4. *In every regular quasi-Markov equilibrium,* $\lim_{\alpha \to 1} s(\alpha, l) > 0$.

Proof of lemma 4: Two preliminary observations: First, if A is willing to recommend option h at belief α even though the bonus option is option l, then necessarily $V(\alpha^+) > 1/\delta$ (as A could get 1 for sure in this period by recommending the bonus option). Second, if $V(\alpha) > 1$, then DM has to play continue (after a fitting recommendation) with positive probability as otherwise $V(\alpha) \leq 1$.

Now let $\mathcal{A} = \{\alpha : \ s(\alpha, l) > 0\}$ and $\bar{a} = \sup \mathcal{A}$. Note that $\mathcal{A}$ is the union of a finite number of intervals by the assumption that the equilibrium strategies are piecewise continuous. If $\bar{a}$ was below 1, then $s(\alpha, l) = 0$ for $\alpha > \bar{a}$ which would imply that DM's best response at these α is to always end the game and therefore $V(\alpha) = 1$ and $W(\alpha) = W_O$ for all $\alpha > \bar{a}$.

I want to show that $\bar{a} = 1$. Note that $\bar{a} = 1$ holds if $\lim_{\alpha \nearrow \bar{a}} s(\alpha, l) > 0$: As $\lim_{\alpha \nearrow \bar{a}} s(\alpha, l) > 0$ implies that $s(\alpha, l) > \varepsilon'$ on $(\bar{a} - \varepsilon, \bar{a})$ for some ε, $\varepsilon' > 0$ by the assumption that s is piecewise continuous, it follows that $\alpha^+ > \bar{a}$ for $\alpha < \bar{a}$ sufficiently close to $\bar{a}$. Following the first preliminary observation above this requires $V(\alpha^+) > 1/\delta$ which contradicts that $V(\alpha) = 1$ for all $\alpha > \bar{a}$. Note that this argument shows more generally that $\alpha^+ \leq \bar{a}$ for all $\alpha \in \mathcal{A}$ (with strict inequality if $\bar{a} < 1$ or $\alpha < \bar{a}$).

Consequently, I only have to rule out $\lim_{\alpha \nearrow \bar{a}} s(\alpha, l) = 0$ to show that $\bar{a} = 1$. To do so, I will show that DM wants to end the game for some $\alpha < \bar{a}$ arbitrarily close to $\bar{a}$ if $\lim_{\alpha \nearrow \bar{a}} s(\alpha, l) = 0$ which contradicts that $s(\alpha, l) > 0$ is a best response for these beliefs. Suppose $\lim_{\alpha \nearrow \bar{a}} s(\alpha, l) = 0$. As the equilibrium is assumed to be piecewise continuous, $s(\alpha, l)$ is continuous on $(\bar{a} - \varepsilon, \bar{a})$ for some $\varepsilon > 0$. Hence, for every $\varepsilon'' > 0$, there exists an $\varepsilon' > 0$ such that $s(\alpha, l) < \varepsilon''$ for $\alpha \in (\bar{a} - \varepsilon', \bar{a})$. Note furthermore that $\lim_{\alpha \nearrow \bar{a}} \alpha - \alpha^- = 0$ as competent and incompetent type use (in the limit) the same strategy as $\alpha \nearrow \bar{a}$. Consequently, for every $T \in \mathbb{N}$ and every $\varepsilon' > 0$, there exists an $\alpha_{T\varepsilon'} \in (\bar{a} - \varepsilon', \bar{a})$ such that, starting from belief $\alpha_{T\varepsilon'}$, the updated belief after T consecutive recommendation that did not fit DM's needs will still be above $\bar{a} - \varepsilon'$. Starting from belief $\alpha_{T\varepsilon'}$, the updated belief of the next T periods will therefore be in $(\bar{a} - \varepsilon', \bar{a})$ as the previous paragraph established that $\alpha^+ \leq \bar{a}$ for all $\alpha \in \mathcal{A}$. Choosing ε'

such that (i) $s(\alpha, l) < \varepsilon''$ for $\alpha \in (\bar{a} - \varepsilon', \bar{a})$ and (ii) $(\bar{a} - \varepsilon', \bar{a}) \subset \mathcal{A}$ yields the following upper bound on $W(\alpha_{T\varepsilon'})$ (recall that by the definition of $\mathcal{A}$ continuing is a best response for all $\alpha \in \mathcal{A}$):

$$\delta^{T+1} \frac{p}{1-\delta} + \sum_{t=0}^{T} \delta^t \left[(1-\alpha)\frac{1}{2} + \alpha \left(p(\frac{1}{2} + \varepsilon'' \frac{1}{2}) + (1-p)\frac{1}{2} \right) \right].$$

This upper bound is strictly less than the W_O by assumption (1) for ε'' sufficiently small and T sufficiently high. Hence, ending the game is DM's unique best response at $\alpha_{T\varepsilon'}$ for T high enough and $\varepsilon' > 0$ small enough which contradicts that $\alpha_{T\varepsilon'} \in \mathcal{A}$. □

Lemma 5. *(5) holds in every regular equilibrium.*

Proof of lemma 5: By piecewise continuity of the equilibrium strategies, there exists an $\bar{a} < 1$ such that V is continuous for all $\alpha \in (\bar{a}, 1)$. By lemma 4 and piecewise continuity, $\bar{a}$ can be chosen high enough to ensure $s(\alpha, l) > 0$ for $\alpha \in (\bar{a}, 1)$. By theorem 1 and bounded variation of β^+ and β^-, $\bar{a}$ can be chosen high enough to also ensure $\min\{\beta^-(\alpha), \beta^+(\alpha)\} < 1$ for $\alpha \in (\bar{a}, 1)$.[19] Note that firing the adviser, i.e. $\beta^-(\alpha) < 1$ or $\beta^+(\alpha) < 1$, is only a best response for high α if $s(\alpha, l) < 1$ (by assumption (1)). Hence, there exists an $\bar{a} < 1$ such that (i) $0 < s(\alpha, l) < 1$ for all $\alpha \in (\bar{a}, 1)$, (ii) V is continuous on $\alpha \in (\bar{a}, 1)$ and (iii) β^+ and β^- are continuous on $\alpha \in (\bar{a}, 1)$. In the remainder of this proof, only beliefs in $\alpha \in (\bar{a}, 1)$ are considered, i.e. (i), (ii) and (iii) are assumed to hold, and the qualifier "for $\alpha \in (\bar{a}, 1)$" is omitted.

As A uses a mixed strategy, he has to be indifferent between recommending either of the two options if option l is the bonus option. That is, the indifference condition $\beta^+(\alpha)V(\alpha^+) - \beta^-(\alpha)V(\alpha^-) = 1/(\delta(2p-1))$ holds; see lemma 2.

As A is indifferent in case option l is the bonus option (and always recommends the bonus option if it is option h), one way to achieve his expected value $V(\alpha)$ is to always recommend his bonus option. Since the bonus option fits DM's needs with probability 1/2, this gives

$$V(\alpha) = 1 + \delta \left(\frac{1}{2}\beta^+(\alpha)V(\alpha^+) + \frac{1}{2}\beta^-(\alpha)V(\alpha^-) \right). \tag{11}$$

Using the indifference condition to eliminate $\beta^+(\alpha)V(\alpha^+)$ gives

$$V(\alpha) = 1 + \frac{1}{2(2p-1)} + \delta\beta^-(\alpha)V(\alpha^-). \tag{12}$$

[19] In detail: Theorem 1 states that there is a sequence of beliefs strictly less than 1 but converging to 1 such that $\min\{\beta^-(\alpha), \beta^+(\alpha)\} < 1 - \varepsilon$ for $\varepsilon = (\eta - \eta\delta)/(2 - 2\eta\delta) > 0$. If there was a sequence of beliefs converging to 1 such that $\min\{\beta^-(\alpha), \beta^+(\alpha)\} = 1$, DM's strategy would exhibit unbounded variation.

V is continuous and bounded; e.g. bounded from above by $1/(1-\delta)$, and bounded from below by 0. Therefore, $\mathring{V} \equiv \lim_{\alpha \to 1} V(\alpha)$ exists. Furthermore, $\lim_{\alpha \to 1} \alpha^- = \alpha$ because updating stops as the belief is approaching 1. The continuity of V implies therefore that $\lim_{\alpha \to 1} V(\alpha) = \lim_{\alpha \to 1} V(\alpha^-)$. Strategies are piecewise continuous by assumption and therefore $\lim_{\alpha \to 1} \beta^-(\alpha)$ exists and will – with a slight abuse of notation – be denoted as $\beta^-(1)$. Taking limits on both sides of (12) yields (after rearranging)

$$\mathring{V} = \frac{4p-1}{(4p-2)(1-\delta\beta^-(1))}.$$

As $\beta^-(1)$ is a probability and therefore in $[0, 1]$, the previous equation yields

$$\mathring{V} \in \left[\frac{4p-1}{4p-2}, \frac{4p-1}{(4p-2)(1-\delta)}\right]. \tag{13}$$

In equation (11), one could also use the indifference condition to eliminate $\beta^-(\alpha)V(\alpha^-)$ (instead of $\beta^+(\alpha)V(\alpha^+)$). This yields

$$V(\alpha) = 1 - \frac{1}{2(2p-1)} + \delta\beta^+(\alpha)V(\alpha^+).$$

Taking again limits on both sides results (after rearranging) in

$$\mathring{V} = \frac{4p-3}{(4p-2)(1-\delta\beta^+(1))}.$$

As $\beta^+(1)$ is a probability and therefore in $[0, 1]$, this implies that

$$\mathring{V} \in \left[\frac{4p-3}{4p-2}, \frac{4p-3}{(4p-2)(1-\delta)}\right]. \tag{14}$$

Taking (13) and (14) together yields the result. $\square$

Finally, note that the upper bound of the interval in (5) is higher than the lower bound of the interval if and only if (4) holds. That is, if (4) does not hold, $\mathring{V}$ cannot both satisfy (13) and (14) at the same time and a regular quasi-Markov equilibrium cannot exist.

The only remaining result is that $W(\alpha) = W_O$ for sufficiently high α. As mentioned in the proof of lemma 5, theorem 1 and bounded variation of β^+ and β^- imply that $\min\{\beta^+(\alpha), \beta^-(\alpha)\} < 1$ for $\alpha \in (\bar{a}, 1)$ for some $\bar{a} < 1$. This implies the result because DM stops the game with positive probability only if $W(\alpha) = W_O$. $\square$

Proof of corollary 1: If $s(\alpha, l) = 0$ for all reputations, then DM best responds by stopping $\beta^+ = \beta^- = 0$ and the result will clearly hold. If $s(\alpha, l)$ is strictly positive on some interval, then, by (5), $\lim_{\alpha\to 1} V(\alpha) \leq (4p-3)/((4p-2)(1-\delta))$. Rearranging

$$\frac{4p-3}{(4p-2)(1-\delta)} < \frac{1}{2(1-\delta)}$$

yields – given that $p > 1/2$ – $p < 1$, which is true by assumption. $\square$

Proof of proposition 4: First, I show that the strategies are feasible under the conditions given in the proposition. Note that $p \geq 1/4 + 1/(2\delta)$ and $\delta < 1$ imply together that $p > 3/4$. Furthermore, it is equivalent to $1-\delta \leq \delta(2p-3/2)$. Hence, β^+ is between 0 and 1. Similarly, $p \geq 1/4 + 1/(2\delta)$ is equivalent to $2p\delta - \delta/2 - 1 \geq 0$ which ensures $\beta^- \geq 0$. Note that $p \leq 1/4 + 1/(2\delta^2)$ is equivalent to $2p\delta - \delta/2 - 1 \leq \delta(1-\delta)(2p-1/2)$ and therefore ensures $\beta^- \leq 1$. The condition $\alpha_1 \geq 2((1-\delta)W_O - 1)/(2p-1)$ leads to $s(\alpha, l) \leq 1$ (while (1) ensures $s(\alpha, l) \geq 0$).

Second, I show how Bayesian updating works out under the proposed strategies. For $\alpha < \alpha_1$, both types act identically and therefore no updating occurs. Generally, the updated beliefs when A uses strategy $s(\alpha, l)$ (denoted by s for short) are

$$\begin{aligned}
\alpha^- &= \frac{\alpha((1-p)/2 + s(1-p)/2 + (1-s)p/2)}{(1-\alpha)/2 + \alpha((1-p)/2 + s(1-p)/2 + (1-s)p/2)} \\
\alpha^+ &= \frac{\alpha(p/2 + sp/2 + (1-s)(1-p)/2)}{(1-\alpha)/2 + \alpha(p/2 + sp/2 + (1-s)(1-p)/2)}.
\end{aligned}$$

With $s(\alpha, l) = (2(1-\delta)W_O - 1)/(\alpha(2p-1))$, this becomes

$$\begin{aligned}
\alpha^- &= \frac{\alpha + 1 - 2(1-\delta)W_O}{2 - 2(1-\delta)W_O} \\
\alpha^+ &= \frac{2(1-\delta)W_O - 1 + \alpha}{2(1-\delta)W_O}.
\end{aligned}$$

Note that $\alpha_1 = \alpha_2^-$ and $\alpha_2 = \alpha_3^-$. Furthermore, $\tilde{\alpha}^+ = \alpha_2$. As the updating rule is monotone, this implies that after a fitting recommendation reputations $\alpha \in [\alpha_1, \tilde{\alpha})$ are updated to $\alpha^+ \in (\alpha, \alpha_2)$ and reputations $\alpha \in [\tilde{\alpha}, \alpha_2]$ are updated to $\alpha^+ \geq \alpha_2$. In the other direction, reputations $\alpha \in [\alpha_1, \alpha_2)$ are updated to $\alpha^- < \alpha_1$ after a non-fitting recommendation and $\alpha \in [\alpha_2, \alpha_3)$ leads to $\alpha^- \in [\alpha_1, \alpha_2)$ while $\alpha \geq \alpha_3$ leads to $\alpha^- \geq \alpha_2$.

Third, note that A's strategy $s(\alpha, l)$ is chosen exactly such that $W(\alpha) = W_O$ for $\alpha \geq \alpha_1$ and therefore DM is indifferent between continuing and stopping whenever his updated belief is weakly above α_1. Clearly, firing the adviser if his updated belief is below α_1 is optimal as the adviser only recommends his bonus option then and therefore DM's value is less than W_O by (1). This implies that DM's strategy is a best response to the adviser's strategy.

Fourth, DM's strategy is chosen such that A is indifferent between recommending either option if the bonus option is l for $\alpha \geq \alpha_1$. To illustrate this, I will go through the different subintervals one by one always assuming that the bonus option is option l. Before doing so, note that A's indifference condition is $1 + p\beta^-(\alpha)\delta V(\alpha^-) + (1 - p)\beta^+(\alpha)\delta V(\alpha^+) = p\beta^+(\alpha)\delta V(\alpha^+) + (1-p)\beta^-(\alpha)\delta V(\alpha^-)$ which can be rearranged to

$$\beta^+(\alpha)V(\alpha^+) - \beta^- V(\alpha^-) = 1/(\delta(2p-1)). \tag{15}$$

Now turn to the first subinterval: If $\alpha \in [\alpha_1, \tilde{\alpha})$, then $\alpha^+ \in (\alpha, \alpha_2)$ and $V(\alpha^+) = (2p - 1/2)/(2p-1)$ while $\alpha^- < \alpha_1$ and therefore $\beta^-(\alpha) = 0$. With these values (15) holds. Next consider $\alpha \in [\tilde{\alpha}, \alpha_2)$. Then $\alpha^+ \geq \alpha_2$ and therefore $V(\alpha^+) = (2p-3/2)/((1-\delta)(2p-1))$ while $\alpha^- < \alpha_1$ and therefore $\beta^-(\alpha) = 0$. With these values (15) holds again. For $\alpha \in [\alpha_2, \alpha_3)$, $\alpha^- \in [\alpha_1, \alpha_2)$ and therefore $V^-(\alpha) = (2p - 1/2)/(2p - 1)$ while $\beta^+(\alpha)V(\alpha^+) = (2p - 3/2)/((1-\delta)(2p-1))$ as $\alpha^+ \geq \alpha \geq \alpha_2$. Hence, the left hand side of (15) becomes

$$\frac{2p - 3/2}{(1-\delta)(2p-1)} - \frac{2p\delta - \delta/2 - 1}{\delta(1-\delta)(2p-1/2)} \frac{2p-1/2}{2p-1} = \frac{\delta(2p-3/2) - (2p\delta - \delta/2 - 1)}{\delta(1-\delta)(2p-1)}$$

which – after canceling – equals the right hand side of (15). Finally, $\alpha^- \geq \alpha_2$ for $\alpha \geq \alpha_3$ and therefore $V(\alpha^-) = (2p - 3/2)/((1-\delta)(2p-1))$ while $\beta^+(\alpha)V(\alpha^+) = (2p - 3/2)/((1-\delta)(2p-1))$. This leads to the same left hand side of (15) as just computed and therefore (15) holds also for $\alpha \geq \alpha_3$. Clearly, A's strategy is optimal for $\alpha < \alpha_1$ and therefore A's strategy is a best response to DM's strategy.

Fifth, the value functions are consistent with the strategies, i.e. the Bellman equations hold. To see this for W, note that A's expected payoff in this period when entering with belief α when A uses strategy s equals $(1-\alpha)/2 + \alpha\,[p/2 + ps/2 + (1-p)(1-s)/2] = (1 - s\alpha(2p-1))/2$. With $s = (2(1-\delta)W_O - 1)/(\alpha(2p-1))$ this is equal to $(1-\delta)W_O$. As A's value in future periods is always δW_O under the strategies above (either because he stops the game or because $W(\alpha) = W_O$), this shows $W(\alpha) = W_O$ for $\alpha \geq \alpha_1$.

Turning to V, recall that DM's strategy is such that A is indifferent between recommending his bonus option or option h when the bonus option is l (for all $\alpha \geq \alpha_1$). Consequently, his expected discounted payoff stream when entering a period with reputation α can be written as if he always recommended option h which would then yield the value $1/2 + p\beta^+(\alpha)\delta V(\alpha^+) + (1-p)\beta^-(\alpha)\delta V(\alpha^-)$. Using the indifference condition (15), this can be written as $1/2 + \beta^+(\alpha^+)\delta V(\alpha^+) - (1-p)/(2p-1) = \beta^+(\alpha^+)\delta V(\alpha^+) + (2p - 3/2)/(2p-1)$. It is straightforward to verify that this equals $V(\alpha)$ for all $\alpha \geq \alpha_1$.

Finally, for $p = 1/4 + 1/(2\delta^2)$ and $\alpha \in [\alpha_2, \alpha_3)$,

$$\beta^-(\alpha) = \frac{2p\delta - \delta/2 - 1}{\delta(1-\delta)(2p-1/2)} = \frac{2\delta/4 + 1/\delta - \delta/2 - 1}{\delta(1-\delta)(1/2 + 1/\delta^2 - 1/2)} = \frac{1/\delta - 1}{(1-\delta)/\delta} = 1.$$

$\square$

References

Aghion, P. and M. Jackson (2016). Inducing leaders to take risky decisions: dismissal, tenure, and term limits. *American Economic Journal: Microeconomics 8*(3), 1–38.

Benabou, R. and G. Laroque (1992). Using privileged information to manipulate markets: Insiders, gurus, and credibility. *Quarterly Journal of Economics 107*(3), 921–958.

Brandenburger, A. and B. Polak (1996). When managers cover their posteriors: Making the decisions the market wants to see. *RAND Journal of Economics 27*(3), 523–541.

Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica 50*(6), 1431–1451.

Cripps, M. W., G. J. Mailath, and L. Samuelson (2004). Imperfect monitoring and impermanent reputations. *Econometrica 72*(2), 407–432.

Fama, E. F. (1980). Agency problems and the theory of the firm. *Journal of Political Economy 88*(2), 288–307.

Fang, L. and A. Yasuda (2009). The effectiveness of reputation as a disciplinary mechanism in sell-side research. *Review of Financial Studies 22*(9), 3735–3777.

Holmström, B. (1982). Managerial incentive problems: A dynamic perspective. Reprinted in *Review of Economic Studies* (1999) *66*(1), 169–182.

Jackson, A. R. (2005). Trade generation, reputation, and sell-side analysts. *Journal of Finance 60*(2), 673–717.

Jullien, B. and I.-U. Park (2014). New, like new, or very good? reputation and credibility. *Review of Economic Studies 81*(4), 1543–1574.

Krishna, V. and J. Morgan (2008). Cheap talk. In S. N. Durlauf and L. E. Blume (Eds.), *New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan.

Mailath, G. J. and L. Samuelson (2001). Who wants a good reputation? *Review of Economic Studies 68*(2), 415–441.

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.

Park, I.-U. (2005). Cheap talk referrals of differentiated experts in repeated relationship. *RAND Journal of Economics 36*(2), 391–411.

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.

Schottmüller, C. (2016). Too good to be truthful: Why competent advisers are fired. TILEC Discussion Paper No. 2016-007.

Sobel, J. (1985). A theory of credibility. *Review of Economic Studies 52*(4), 557–573.

Wiseman, T. (2008). Reputation and impermanent types. *Games and Economic Behavior 62*(1), 190–210.