Supplementary material to
Too good to be truthful: Why competent advisers are fired

Christoph Schottmüller*
University of Cologne

February 28, 2019

1. Relaxing assumptions

I will restate theorems 1 and 2 in the paper (and give somewhat different proofs) for a more general model. The results are qualitatively the same although the exact expressions for the bounds change.

The differences to the model in the paper are the following. First, A and DM have different discount factors $\delta_A \in (0,1)$ and $\delta_D \in (0,1)$.

Second, the signal technology is generalized in several ways. I allow a general history dependent and time varying signal technology. To be more precise, note that the signal technology can be summarized by a competent type's posterior belief, i.e. the probability that he assigns to option h fitting DM's needs. I denote by p_{tiH} this probability in period t when the bonus option is option $i \in \{l, h\}$ and H is the history of the game in periods 0 to $t-1$. I assume that all $1/2 < p_{tiH} < 1$ for all $t = 0, 1, \dots$, all $i \in \{l, h\}$ and possible histories H. More precisely, I make the slightly stronger assumption that the supremum of all the p_{tiH}, denoted by $\bar{p}$, is strictly below 1 and the infimum of all the p_{tiH}, denoted by $\underline{p}$, is strictly above 1/2. Apart from these assumptions, I do not restrict the signal technology further.

Third, the information structure can be generalized as long as the assumption is maintained that DM cannot tell the options apart. That is, DM cannot infer from the identity of a recommendation whether it is more or less likely to be the bonus option. I will ensure this by maintaining the assumption that each option is equally likely to fit DM's needs and equally likely to be the bonus option. Under this condition allowing for imperfect and time varying correlation will not affect the analysis: Call the two

*email: c.schottmueller@uni-koeln.de

options *option 1* and *option 2*. Let $\kappa_{tij} \in (0,1)$ for $t = 0, 1, \dots$, $i \in \{f, n\}$, $j \in \{b, n\}$ denote the probability that option 1 in period t fits ($i = f$) or does not fit DM's needs while being the bonus option ($j = b$) or not. Assume that the infimum of all κ_{tij} is strictly above 0 and the supremum strictly below 1. The assumption that both options are equally likely to fit DM's needs and equally likely to be the bonus option require then $\kappa_{tfb} + \kappa_{tfn} = 1/2$ and $\kappa_{tfb} + \kappa_{tnb} = 1/2$. In addition, $\kappa_{tfb} + \kappa_{tfn} + \kappa_{tnb} + \kappa_{tnn} = 1$. Hence, for a given t all κ_{tij} are determined if we know the probability that the decision maker's and the adviser's interests coincide. Call this parameter $\kappa_t \equiv \kappa_{tfb} + \kappa_{tnn}$. Denote by $\bar{\kappa}$ the supremum of all κ_t and by $\underline{\kappa}$ the infimum. In the paper, $\kappa_t = 1/2$ for all periods while now different values are allowed. However, it should be admitted that the possibility of this generalization is not very surprising given the assumption below that states that advice by the incompetent type is still worse than the outside option: For A, the whole correlation structure is immaterial given his posterior p_{tiH} because he has to make a decision when his belief is described by this posterior and the prior belief structure as described by the κ_{tij} is at this point simply no longer relevant. For DM the correlation is also not particularly relevant as long as (i) he cannot draw inference from the identity of the recommendation and (ii) in every period he prefers the outside option to incompetent advice as I will assume.

The parameter assumptions from the paper have to be adapted to the new setting: I assume that $\bar{\kappa} < (1 - \delta_D)W_O < \underline{p}$, i.e. the per period payoff from the outside option is more attractive than advice from an incompetent type in this period but worse than best possible advice from a competent type.[1] Furthermore, I assume $p_{tiH} - \kappa_t$ is bounded from below by a strictly positive number. Note that $p_{tiH} > \kappa_t$ simply states that a competent type's signal is informative (otherwise A would give better advice by recommending his bonus option than by recommending option h). That the difference is greater than some strictly positive number is a slight strenghening of this condition. Also recall the earlier stated parameter assumptions $\underline{\kappa} > 0$ and $\bar{\kappa} < 1$ and $\bar{p} < 1$.

Now it is possible to generalize theorem 1.

Theorem 3. *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 $\varepsilon = (1 - \delta_A)^2/(2\delta_A^2) > 0$ after all histories after which his reputation α is greater than $\bar{\alpha}$.*

Proof of theorem 3: Suppose the statement was not true; i.e. suppose that there was an equilibrium and an $\bar{\alpha} < 1$ such that DM ends the game with probability ε or less after every history leading to a belief above $\bar{\alpha}$.

[1] The proofs below go through without change if one assumes the weaker condition $\bar{\kappa} < (1-\delta_D)W_O < \bar{p}$ which allows for a somewhat higher outside option. This is not surprising as a higher outside option will make it more attractive to take this outside option.

Note that Bayesian updating becomes arbitrarily slow for α close to 1.[2] Hence, for every $T \in \mathbb{N}$ there exists an $\hat{\alpha}_T \in (\bar{\alpha}, 1)$ such that A's belief after T consecutive non-fitting recommendations will still be above $\bar{\alpha}$ whenever the initial α is above $\hat{\alpha}_T$.

I will now show that this implies that A will always recommend his bonus option if $\alpha > \hat{\alpha}_T$ for some sufficiently high T. Clearly, A will recommend his bonus option if it is option h. Hence, let us assume that A's bonus option is option l and that $\alpha > \hat{\alpha}_T$. Note that A can achieve an expected payoff of at least $1+\delta_A(1-\varepsilon)+\delta_A^2(1-\varepsilon)^2+\cdots+\delta_A^T(1-\varepsilon)^T$ by always recommending his bonus option. If A does not recommend his bonus option in the current period, an upper bound on his expected payoff is $\delta_A/(1-\delta_A)$ (which is the payoff from recommending option h initially and the bonus option ever after under the additional assumption that DM never stops the game). I will now show that the lower bound on A's payoff from recommending his bonus option is higher than the upper bound on his payoff from recommending option l (for sufficiently high T):

$$\begin{aligned}
1+\delta_A(1-\varepsilon)+\delta_A^2(1-\varepsilon)^2+\cdots+\delta_A^T(1-\varepsilon)^T &> \frac{\delta_A}{1-\delta_A} \\
\Leftrightarrow \frac{1-\delta_A^T(1-\varepsilon)^T}{1-\delta_A(1-\varepsilon)} &> \frac{\delta_A}{1-\delta_A} \\
\Leftrightarrow 1-\delta_A-\delta_A^T(1-\varepsilon)^T(1-\delta_A) &> \delta_A-\delta_A^2(1-\varepsilon) \\
\Leftrightarrow \varepsilon &< 1-\frac{2\delta_A-1}{\delta_A^2}-\delta_A^{T-2}(1-\varepsilon)^T \\
\Leftrightarrow \varepsilon &< \frac{(1-\delta_A)^2}{\delta_A^2}-\delta_A^{T-2}(1-\varepsilon)^T.
\end{aligned}$$

The last term tends to zero as T becomes arbitrarily high. Given the value of ε in the theorem, the inequality is therefore true for high enough T. Formally, let T be the lowest natural number such that $2\delta_A^{T-2}(1-((1-\delta_A)/(2\delta_A))^2)^T < ((1-\delta_A)/(\delta_A))^2$. Then A will always recommend his bonus option if $\alpha > \hat{\alpha}_T$.

This, however, contradicts that DM plays best response: For every $\alpha > \hat{\alpha}_T$, both the competent and the incompetent type use the same strategy – always recommending the bonus option. This implies that after every history H leading to a belief $\alpha > \hat{\alpha}_T$, DM will not update his belief in the period after H and therefore his belief in the next period (if he continues) will still be $\alpha > \hat{\alpha}_T$. As we just established that in this case A recommends his bonus option, DM's belief will also not update then. Iterating the argument yields that $\alpha > \hat{\alpha}_T$ in all future periods (if DM continues) and A will recommend his bonus option in all future periods. Hence, DM's expected payoff after a history leading to belief $\alpha > \hat{\alpha}_T$ is at most $\bar{\kappa}/(1-\delta_D)$ if he continues. This upper

[2] More precisely, α^- for a given α is bounded from below by $(\alpha(1-\bar{p}))/(\alpha(1-\bar{p})+(1-\alpha)(1-\underline{\kappa}))$. Consequently, $\alpha-\alpha^-$ is bounded from above by $\alpha-(\alpha(1-\bar{p}))/(\alpha(1-\bar{p})+(1-\alpha)(1-\underline{\kappa})) = \alpha\frac{(1-\alpha)(\bar{p}-\underline{\kappa})}{\alpha(1-\bar{p})+(1-\alpha)(1-\underline{\kappa})}$ which converges to zero as $\alpha \to 1$.

bound on his payoff is less than W_O by assumption. Hence, continuing with probability of at least $1-\varepsilon > 0$ is not a best response which is the desired contradiction. $\square$

To obtain an equivalent to theorem 2, the following lemma is a useful intermediate step.

Lemma 6. *Let $\varepsilon > 0$ and define*

$$T_\varepsilon = \left\lceil \frac{log(\varepsilon)}{log(1-(1-\bar{p})^{T'+T''}\varepsilon')} \right\rceil \qquad where$$

$$\varepsilon' = \frac{1-\delta_A}{2} \quad and \quad T' = \left\lceil 2\frac{log(1-\delta_A)}{log(\delta_A)} - 1 \right\rceil \quad and \quad T'' = \left\lceil \frac{log\left(\frac{(1-\delta_D)W_O-\bar{\kappa}}{\delta_D(\bar{p}-\bar{\kappa})}\right)}{log(\delta_D)} \right\rceil.$$

In equilibrium, the probability that DM ends the game within T_ε periods is at least $1-\varepsilon$ after every history.

Proof of lemma 6: Take an arbitrary history up to period t as given and keep it fixed for the rest of the proof.

First, I want to establish that A will always recommend his bonus option if DM continues for the next T' periods with at least probability $1-\varepsilon'$ on every continuation path of length T'.

By means of contradiction suppose there is no continuation path of length T' such that DM ends the game on this path with probability ε' or higher. I will show that in this case A will have an incentive to recommend his bonus option even if the bonus option is option l. By recommending the bonus option, A can achieve a payoff of at least $1+(1-\varepsilon')(\delta_A+\delta_A^2+\cdots+\delta_A^{T'})$. By not recommending the bonus option A will achieve a payoff of at most $0+\delta_A/(1-\delta_A)$. With ε' and T' chosen as above, however, the lower bound on the payoff from recommending the bonus option is higher than the upper bound on the payoff of recommending option h:

$$\begin{aligned}
T' = \left\lceil 2\frac{\log(1-\delta_A)}{\log(\delta_A)} - 1 \right\rceil &\geq 2\frac{\log(1-\delta_A)}{\log(\delta_A)} - 1 \\
\Leftrightarrow (T'+1)\log(\delta_A) &\leq 2\log(1-\delta_A) \\
\Leftrightarrow \delta_A^{T'+1} &\leq (1-\delta_A)^2 \\
\Leftrightarrow \frac{\delta_A^{T'+1}}{1-\delta_A} &\leq 1-\delta_A \\
\Leftrightarrow \delta_A\left(1+\delta_A^{T'+1}+\delta_A^{T'+2}+\dots\right) &\leq 1 \\
\Leftrightarrow (1-\delta_A)\left(\delta_A+\delta_A^2+\cdots+\delta_A^{T'}\right)+\delta_A^{T'+1}+\delta_A^{T'+2}+\dots &\leq 1 \\
\Leftrightarrow \delta_A+\delta_A^2+\dots &\leq 1+\delta_A\left(\delta_A+\delta_A^2+\cdots+\delta_A^{T'}\right) \\
\Leftrightarrow \frac{\delta_A}{1-\delta_A} &\leq 1+(1-2\varepsilon')\left(\delta_A+\delta_A^2+\cdots+\delta_A^{T'}\right)
\end{aligned}$$

which implies that the lower bound on the payoff of recommending the bonus option (which is the right hand side but without multiplying ε' by 2) is strictly higher than the upper bound on the payoff from recommending option h. Consequently, A will recommend the bonus option even if it is option l in equilibrium (after the initially fixed history). This establishes that ε' and T' have the desired property.

Second, I want to establish that DM optimally ends the game after every history such that according to his equilibrium strategy A will recommend his bonus option regardless of signal in the next T'' periods. Suppose therefore that A will recommend his bonus option in the next T'' periods. Then DM's payoff from continuing is bounded from above by

$$\frac{\bar{\kappa}(1-\delta_D^{T''+1})}{1-\delta_D}+\frac{\delta_D^{T''+1}\bar{p}}{1-\delta_D}.$$

This upper bound is less than W_O as

$$\begin{aligned}\frac{\bar{\kappa}(1-\delta_D^{T''+1})}{1-\delta_D}+\frac{\delta_D^{T''+1}\bar{p}}{1-\delta_D} &< W_O\\ \Leftrightarrow \delta_D^{T''}(\delta_D\bar{p}-\delta_D\bar{\kappa}) &< (1-\delta_D)W_O-\bar{\kappa}\\ \Leftrightarrow T'' &> \frac{log\left(\frac{(1-\delta_D)W_O-\bar{\kappa}}{\delta_D(\bar{p}-\bar{\kappa})}\right)}{log(\delta_D)}\end{aligned}$$

where the last inequality holds true by the definition of T''.

Third, I combine the results from the previous two steps to conclude that there is at least one continuation path of play of length (up to) $T''' = T' + T''$ such that the probability that DM stops the game along this path is at least $\varepsilon' > 0$. By way of contradiction, suppose this was not the case. A would then in the next T'' periods always recommend his bonus option by the first step. By the second step, DM would then find it optimal to stop the game immediately which is the desired contradiction.

As the continuation path of length T''' on which DM's probability to end the game has positive probability under equilibrium play, by the assumption that A is uncertain (i.e. $\bar{p} < 1$), it follows that the game ends with probability $\gamma\varepsilon' > 0$ in the next T''' periods where γ is a lower bound on the probability of the path occurring under equilibrium play which can be chosen independent of the specifics of the equilibrium and the belief, i.e. depending only on the precision of A's signal. For example, $\gamma = (1-\bar{p})^{T'}$ works and will be used in the remainder.

Hence, the probability that DM does not end the game within $2T'''$ periods is at most $(1-\gamma\varepsilon')^2$. Iterating yields that the probability that DM does not end the game within mT''' periods is at most $(1-\gamma\varepsilon')^m$. Let m' be such that $\varepsilon > (1-\gamma\varepsilon')^{m'}$, and let

$T_\varepsilon > m'T'''$. With $\gamma = (1-\bar{p})^{T'''}$ and $T''' = T' + T''$ this implies that

$$T_\varepsilon = \left\lceil \frac{\log(\varepsilon)}{\log\left(1-(1-\bar{p})^{T'+T''}\varepsilon'\right)} \right\rceil (T'+T'')$$

will satisfy the requirements and the lemma follows. □

Now a result analogous to theorem 2 in the paper can be stated. That is, an upper bound on the expected length of the game that is independent of equilibrium and initial belief can be derived.

Theorem 4. *The expected length of the advice relationship in equilbirum is finite and bounded from above by*

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

Proof of theorem 4: Lemma 6 states 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

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

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

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

□

2. Cheap talk

In this section I want to briefly outline in which sense the model of the paper where the adviser's recommendation is directly payoff relevant yields the same results as a model where (i) the adviser's recommendation is cheap talk and (ii) DM is aware of the two options. That is, DM has a choice whether to follow the recommendation or to pick the other option. This implies that one has to be more precise about what the adviser observes. I will keep the assumption that the adviser can observe whether the chosen action fitted DM's needs or not (which implies that the adviser's reputation is common knowledge after every history in a given equilibrium). I will also assume that the adviser can observe whether his recommendation was followed which seems logical as this is payoff relevant to A: A should know whether his bonus option was chosen or not.

A first implication is that all equilibria from the paper are still equilibria in the cheap talk model in the following sense: A strategy in the model of the paper tells a player what to do after every possible history of fitting and non-fitting recommendations. For DM this means that his strategy assigns to every sequence of per period outcomes $(o_1, o_2, \ldots, o_t)$ with $o_i \in \{fit,\ no\ fit\}$ a probability of stopping the game.[3] When reinterpreting DM's strategy for the cheap talk game the following is crucial: o_i denotes whether the recommendation of period i fitted DM's needs and not whether the chosen action fitted his needs! Now take some equilibrium from the model with directly payoff relevant recommendations. Let DM's strategy in the cheap talk game be such that he always follows the adviser's recommendation and uses the same stopping strategy as in the equilibrium with payoff relevant recommendations. Let A's strategy be the same in the cheap talk game as in the game with payoff relevant recommendations. Then these strategies are an equilibrium in the cheap talk game because – by the restriction to *informative* equilibria – DM always finds it weakly optimal to follow the recommendation. To see this, take an arbitrary history H under which the sequence of outcomes $(o_1, o_2, \ldots, o_t)$ realizes in a given equilibrium from the model in the paper. Denote DM's continuation value after the sequence of outcomes $(o_1, o_2, \ldots, o_t)$ and a fitting (non-fitting) recommendation by W^+ (W^-) and let the DM's expected probability of receiving fitting advice after $(o_1, o_2, \ldots, o_t)$ be q. Then following the recommendation yields an expected payoff of $q + q\delta W^+ + (1-q)\delta W^-$ while not following yields an expected payoff of $1 - q + q\delta W^+ + (1-q)\delta W^-$. The crucial point is that whether DM follows or does not follow the recommendation has no impact on the continuation payoffs as strategies are contingent on whether the recommendation fitted DM's needs (and not on whether the chosen option fitted DM's needs). By informativeness of the equilibrium $q \geq 1 - q$ and therefore it is optimal to follow the recommendation. As the stopping and recommendation decisions are the same as in the game with payoff relevant recommendations, neither player has an incentive to deviate in these decision

[3] Note that it would be imprecise to call $(o_1, o_2, \ldots, o_t)$ a "history" as the history has to also include a sequence of A's decisions, i.e. whether to recommend the bonus option or not and what the bonus option was. However, DM's strategy cannot depend on this information as DM does not observe it.

in the cheap talk game. Hence, the reinterpreted equilibrium strategies of the game with payoff relevant recommendations form an equilibrium of the cheap talk game.

The previous arguments also show to which extent the reverse implication is true. If in an equilibrium of the cheap talk game (i) strategies do not depend on the history of chosen actions (but only on the history of fitting or non-fitting recommendations) and (ii) DM does not mix in his decision to follow or not, then there is a payoff equivalent equilibrium in the game with payoff relevant recommendations. This is obvious if DM always follows the recommendation in the cheap talk equilibrium. If after some histories DM does not follow the recommendation, one has to "invert" strategies after these histories first. That is, DM not following A's recommendation if A recommends option h with probability ρ_k when his bonus option is $k \in \{l, h\}$ is equivalent to DM following A's recommendation and A recommending option h with probability $1 - \rho_k$ if his bonus option is $k \in \{l, h\}$. This inversion is essentially a relabeling of the recommendation and does therefore not affect payoffs. It follows that – when focussing on equilibria of the cheap talk game in which the DM's decisions to follow or not to follow are pure – it is without loss of generality to restrict oneself to equilibria where DM always follows the recommendation.

It turns out that the restriction (ii), i.e. the restriction to equilibria in which DM does not mix in his decision to follow or not, is – given restriction (i) – not as restrictive as it might seem. Given restriction (i), DM follows the recommendation whenever the probability that the recommendation fits his needs is strictly greater than 1/2. The reason is the same as stated above: following the recommendation yields an expected payoff of $q + q\delta W^+ + (1 - q)\delta W^-$ while not following yields an expected payoff of $1 - q + q\delta W^+ + (1 - q)\delta W^-$ (same notation as above). Hence, it is optimal to follow the recommendation whenever $q > 1/2$. Using the "strategy inversion" (i.e. relabeling of the recommendations) from the previous paragraph, I can restrict myself without loss of generality to equilibria where $q \geq 1/2$ after every history. When $q = 1/2$ the competent type and the incompetent type necessarily use the same strategy of recommending their bonus option (regardless of whether it is option h or l). Only in this case, DM is indifferent and could potentially use a mixed strategy.

A particular case, are Markov (and quasi-Markov) equilibria. Here restriction (i) is automatically satisfied. Furthermore, restriction (ii) is unnecessary because $q = 1/2$ will not happen on the equilibrium path: $q = 1/2$ after some history implies that competent and incompetent type act in the same way and therefore the belief updating stops. Hence, DM would have an expected continuation payoff below W_O after such a history. Consequently, DM is better off by stopping the game in the previous period. Therefore, the set of (quasi-) Markov equilibria payoff vectors is the same in the cheap talk game and the game with payoff relevant recommendations.

3. Transferable utility

3.1. Efficiency with transfer from adviser to decision maker

Take the setting described in section 4.1 with monetary transfers. I will here derive conditions under which full efficiency can be reached if transfers from A to DM are feasible. To this end it will be unneccessary to have transfers from DM to A and I will therefore ignore these. With respect to timing it is easiest (though not essential for the qualitative results) to assume that the adviser can make a tranfer payment to DM in every period right after the players observe whether the advice fitted DM's needs and before DM decides whether to continue or stop the game.

Consider the following strategies: The adviser (regardless of type) always recommends the bonus option and pays $\eta = (1-\delta)W_O - 1/2$ to DM whenver this period's recommendation did not fit DM's needs. DM continues in a given period if either he received fitting advice or he received a transfer of at least η. Off path, i.e. if DM continued although re received less than η and bad advice in some period, A always recommends his bonus option and does not make any payment to DM and DM always stops the game.

It remains to check that no player has a profitable deviation. As the proposed strategies are stationary, DM's expected discounted payoff stream at the start of every period is the same on the equilibrium path: $W = 1/2 + 1/2\eta + \delta W$ which can be solved for $W = W_O$. Hence, DM is always indifferent between continuing and stopping on the equilibrium path and therefore his strategy is a best response. Off the equilibriu path, stopping is clearly a best response as the outside option is better than being recommended the bonus option forever by assumption. Next turn to the incompetent type. The value of his expected discounted payoff stream in this equilibrium on path is given by $V = 1 - \eta/2 + \delta V$ which can be solved for $V = (5/4)/(1-\delta) - W_O$ which is strictly positive by $W_O < p/(1-\delta)$. Clearly recommending his bonus option is optimal for the incompetent type in every period. Paying η to DM after non-fitting advice leads to an expected discounted payoff stream of $-\eta + \delta V = (5\delta/4)/(1-\delta) - \delta W_O - (1-\delta)W_O + 1/2$ which is weakly greater than 0 if and only if $\delta \geq (W_O - 1/2)/(3/4 + W_O)$. As a deviation to a lower payment to DM would lead to the end of the game and therefore 0 payoffs, the incompetent type's strategy is a best response if and only if $\delta \geq (W_O - 1/2)/(3/4 + W_O)$ (deviating to higher payments than η is clearly not a best response). Finally, consider the competent type. Note that the value of his expected discounted payoff stream on path is also V and that for the same argument as for the incompetent type paying η in case of non-fitting advice is a best response if and only if $\delta \geq (W_O - 1/2)/(3/4 + W_O)$. We still need to check whether the competent type can profitably deviate by recommending option h instead of the bonus option (recommending option l is clearly not a profitable deviation). This is, of course, only a deviation if option h is not the bonus option. In this case, recommending option h yields an expected payoff of $p\delta V + (1-p)(-\eta + \delta V) = -(1-p)\eta + \delta V$ while recommending the bonus option yields a payoff of $1 + p(-\eta + \delta V) + (1-p)\delta V = 1 - p\eta$. The latter exceeds the former if $p \leq (\eta + 1)/(2\eta)$.

Therefore, the strategy profile above is an equilibrium if (i) $\delta \geq (W_O - 1/2)/(3/4 + W_O)$ and (ii) $p \leq (\eta + 1)/(2\eta)$.

Note that the equilibrium above is not really about advice: In fact, DM is paid to do something he would otherwise not do (follow A's advice perpetually though A is recommending his bonus option all the time). However it is easy to adapt the equilibrium above to feature meaningful advice. Consider the case where (ii) is not satisfied, i.e. $p > (\eta + 1)/(2\eta)$. Following the same steps as above, it is then straightforward to verify that there is an equilibrium in which the incompetent type and DM use the same strategies as above and the competent type is always recommending option h (of course again provided that (i) holds). In this equilibrium, DM even has a rent since the advice he receives from the competent type is informative.[4] In conclusion, there exists an efficient equilibrium if players are sufficiently patient, i.e. if $\delta \geq (W_O - 1/2)/(3/4 + W_O)$.

3.2. Scaling the adviser's bonus

Following the second caveat in section 4.1 of the paper, assume that the adviser's payoff from recommending the bonus option is κ instead of 1. Here I consider transfers from DM to A only which can be paid at the same time at which DM decides whether to continue. The goal is to show that for sufficiently low κ there is an equilibrium in which DM continues for sure if α is above a certain threshold.

To this end, consider the following strategies: On the equilibrium path, DM continues if the updated belief α is weakly above

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

where $w = \kappa/(2p-1)$ and stops the game otherwise. If DM receives fitting advice and continues the game, he also makes a payment of $w = \kappa/(2p-1)$ to A. Off the equilibrium path, DM stops the game and makes no payment. The incompetent type always recommends his bonus option. The competent type always recommends option h on the equilibrium path and always the bonus option off the equilibrium path; i.e. if either DM continued although $\alpha < \underline{\alpha}$ or if DM did not pay w after fitting advice (and strictly speaking also in case DM made a payment although the advice did not fit but this could be changed without affecting what follows) A will switch to "babbing".

[4]Note that DM nevertheless wants to stop the game in case he receives a payment less than η after non-fitting advice: The reason is that this is an off path action and the way the equilibrium was constructed babbling without payment ensues off path.

To make this strategy profile interesting, assume that $\alpha_0 > \underline{\alpha}$. Note that $\underline{\alpha} < 1$ for κ sufficiently small by the assumption $(1-\delta)W_O < p$.

Given these strategies, I will denote A's expected discounted payoff stream at the start of a period by $V(\alpha)$. Clearly, V is weakly increasing: Take any arbitrary sequence of nature's future moves and notice that the discounted payoff stream of A along this sequence is weakly higher if the initial α is higher (under the assumption that players stick to the strategies above). As this is true for any sequence of nature's moves, it is also true in expectation.

The incompetent type's strategy as well as off path play are clearly optimal. The competent type obviously does not want to deviate when the bonus option is option h. If the bonus option is option l (and play is on path) deviating to recommending the bonus option would yield a payoff of $\kappa + (1-p)(w + \delta V(\alpha^+)) + p\delta V(\alpha^-)\mathbb{1}_{\alpha^- \geq \underline{\alpha}}$ while recommending option h yields $p(w + \delta V(\alpha^+)) + (1-p)\delta V(\alpha^-)\mathbb{1}_{\alpha^- \geq \underline{\alpha}}$. Given $w = \kappa/(2p-1)$, the deviation is unprofitable if and only if $V(\alpha^+) \geq V(\alpha^-)\mathbb{1}_{\alpha^- \geq \underline{\alpha}}$ which is true by the monotonicity of V. It follows that A's strategy is indeed a best response. Now turn to DM's strategy. Note that $\underline{\alpha}$ was chosen sufficiently high so that DM indeed prefers continuing (and paying w) to stopping the game (and paying nothing) if the updated belief is above $\underline{\alpha}$: A lower bound on DM's expected payoff stream when continuing and paying w is $-w + \delta[\alpha p + (1-\alpha)/2 + \delta W_O]$ which is larger than the payoff from stopping, i.e. δW_O, by $\alpha \geq \underline{\alpha}$. Given the off path construction, all other deviations are clearly not profitable. Hence, the above is an equilibrium if $\underline{\alpha} < 1$ which is equivalent to $\kappa < \delta(2p-1)(p - 1/2 - \eta)$ where $\eta = (1-\delta)W_O - 1/2$. (By assumption, $\eta \in (0, p-1/2)$ and the cutoff for κ is therefore strictly positive).

4. Competition among advisers: the option of reset

As a third setup – in addition to the two in section 4.2 of the paper – I want to consider the possibility that DM has the options to either *continue* or to *reset*. Reset means that DM starts with a new adviser of a given reputation α_0. That is, the outside option W_O is again endogenous. In contrast to the second setting, DM will not return to advisers he previously fired. A similar line of thoughts as earlier shows the same results as the second setting: A will always recommend his bonus option for sufficiently high α if there is an $\bar{\alpha}$ such that DM continues for sure whenever $\alpha > \bar{\alpha}$. A will always recommend his bonus option if the expected length of the advice relationship with the current adviser exceeds $\bar{T}$ as given in theorem 2. Put differently, analogues to theorems 1 and 2 hold. Furthermore, the following result is similar to the sufficiency part of proposition 3.

Proposition 4. *A regular quasi-Markov equilibrium exists if (4) holds.*

Proof of proposition 4: The idea is to create the equilibrium in the sufficiency part of the proof of proposition 3 where W_O is chosen in such a way that resetting gives indeed this value to the decision maker. To get actual communication $\underline{\alpha}$ should be below α_0. I will construct an equilibrium with $\underline{\alpha} = \alpha_0/2$. By proposition 3, $\underline{\alpha} = (W_O(1-\delta) - 1/2)/(p-1/2)$ has to hold in a regular quasi-Markov equilibrium. Rearranging and using $\underline{\alpha} = \alpha_0/2$ gives

$$W_O = \frac{\alpha_0(p-1/2)+1}{2(1-\delta)}.$$

Consider the equilibrium strategies and value functions as in the sufficiency part of the proof of proposition 3 with this W_O (where V and s denote now the value function and strategy for each adviser). As shown there, these strategies are mutual best responses (given the value of the outside option W_O). As $W(\alpha_0) = W_O$, W_O is indeed DM's value of resetting and therefore these strategies and value functions form an equilibrium of the game with resetting and endogenous outside option. □

5. Bounds on A's payoffs for all perfect Bayesian equilibria

This appendix illustrates that A does not necessarily face a commitment problem in all perfect Bayesian equilibria, i.e. the bounds on A's expected discounted payoff stream in (5) hold only for quasi-Markov equilibria. To show this, I will first derive a relatively simple equilibrium in which the value of DM's expected discounted payoff stream is higher than W_O. This equilibrium will later be used as a continuation after certain histories.

Consider the following strategies: DM continues the game if he received fitting advice in the current and all previous periods and ends the game otherwise. A's strategy is to recommend option h on equilibrium path. Off path, i.e. if DM continued despite receiving non-fitting advice before, A always recommends his bonus option. The value of DM's expected discounted payoff stream equals then $\hat{W}(\alpha) = \alpha p + (1-\alpha)/2 + \delta\Big((\alpha p + (1-\alpha)/2)\hat{W}(\alpha^+) + (1-\alpha p-(1-\alpha)/2)W_O\Big)$ (at the start of the game and also after all histories in which only fitting advice was provided).[5] A's expected discounted payoff stream is valued $\hat{V} = 1/2 + \delta p\hat{V}$ (at the start of the game and at the start of every period in which the game was not stopped yet) which can be solved for $\hat{V} = (1/2)/(1-\delta p)$. It is straightforward to verify that A's strategy of recommending option h is a best response if and only if $p \geq 1/4 + 1/(2\delta)$. DM's strategy is a best response if α_0 is not too low which we will assume in the remainder.[6] Clearly, $W(\alpha_0) > W_O$ in this case as A will give best possible advice as long as the

[5] I use the same notation as for quasi-Markov equilibria here but I should point out that this is clearly not a quasi-Markov equilibrium. In particular, $\hat{W}(\alpha)$ denotes DM's value at histories leading to belief α without containing any non-fitting advice. After advice not fitting his needs, DM ends the game and his value will there only be the outside option.

[6] Note that $\hat{W}(\alpha^+) \geq W_O$ and therefore a simple sufficient condition would be $\alpha_0 > [(1-\delta)W_O - 1/2]/(p-1/2)$.

game continues. Denote the value of DM's expected discounted payoff stream in this equilibrium at the start of the game by $\hat{W}(\alpha_0)$.

Now consider another equilibrium (candidate): For T periods A will recommend his bonus option and DM will continue. In period $T+1$ the two players will start to play the simple equilibrium of the previous paragraph. Under the same conditions – namely p and α_0 sufficiently high – A's strategy is clearly a best response to DM's strategy and DM's strategy is a best response from period $T+1$ onward. It remains to verify that DM cannot profitably deviate before period $T+1$. In particular, I will verify that stopping the game immediately is not a profitable deviation. This deviation is unprofitable if $\delta^T \hat{W}(\alpha_0) + \sum_{t=0}^{T-1} \delta^t/2 \geq W_O$ which is true for W_O sufficiently low. In particular, the inequality holds, due to $\hat{W}(\alpha_0) > W_O$, strictly at the lower bound for W_O which is $(1/2)/(1-\delta))$, see (1). This shows that the above is an equilibrium for a given finite T if W_O is sufficiently low. Note that A's expected discounted payoff stream (from the start of period 1 onward) in this equilibrium equals $(1-\delta^T)/(1-\delta) + \delta^T \hat{V}$. For T sufficiently high, this is clearly higher than $(1/2)/(1-\delta)$ which is the hypothetical commitment payoff that would result from committing to recommending option h each period.

For $T \to \infty$ and $W_O \to (1/2)/(1-\delta)$, the just derived equilibrium payoff in fact attains the upper bound $\bar{V}$ derived in the proof of theorem 1, see (9). Hence, the lowest upper bound on A's equilibrium payoff is $1/(1-\delta)$. The infimum of A's equilibrium payoffs is clearly attained in the babbling equilibrium in which DM stops the game immediately.