Anonymous or personal?
A simple model of repeated personalized advice*

Marius Gramb† — University of Cologne
Christoph Schottmüller‡ — University of Cologne

December 9, 2022

Abstract

A consumer repeatedly asks an expert for advice. The expert's incentives are not aligned with the consumer's preferences because he can receive a bonus if the consumer takes certain actions. Over time, the expert gets to know the consumer and is therefore able to give better advice (if he wants to do so). In simple equilibria, both – consumer and expert – benefit from the expert's learning if "learning" is such that the expert's best guess about what is the best advice for the consumer becomes more precise. This provides a natural explanation for why consumers have a preference for personalized advice and also for why most internet users do not use anonymization tools. The theoretical predictions are tested in a laboratory experiment.

JEL codes: C73, C91, D82, D83

Keywords: Advice, Cheap Talk, Privacy

1. Introduction

In many situations, consumers ask better-informed experts to guide their choices. This happens even in situations where experts may have preferences over consumer choices that do not match the consumers' preferences, and it happens even in situations where it is difficult for the consumer to accurately articulate his exact preferences. For example, a consumer might ask his bank's employees for financial advice. The bank employee

*We gratefully acknowledge the financial support of the Center for Social and Economic Behavior (C-SEB) of the University of Cologne.

†email: mgramb@uni-koeln.de

‡email: c.schottmueller@uni-koeln.de

typically receives a bonus if the consumer purchases a particular investment product and often different investment products result in different bonuses for the adviser. There is no reason to believe that the product with the highest bonus is also the one best suited for the consumer. Similar situations occur in other retail sectors, such as consumer electronics or even cars.

Another example is internet search. A consumer enters a search term and relies on the search engine's response. Since some links are sponsored, there is an incentive for the search engine to emphasize the sponsored links more than links that better fit the consumer's needs but are not sponsored. A third example would be a minister (or manager) asking a civil servant (subordinate) to draft a particular legislative act or decree. Even if the civil servant has no policy preferences of his own, he might be aware that a similar draft has already been written under a previous government and that handing that old draft to the minister would save him a lot of time and effort. Again, this old draft is unlikely to do exactly what the minister wanted to accomplish. As a final example, consider a physician-patient relationship. The patient describes his symptoms and the physician prescribes a medication. Given the lobbying efforts of the pharmaceutical companies, it is quite possible that the physician has a preference for a certain drug company or pharmaceutical product.

What do these examples have in common? A consumer asks an expert to help him make a choice, although he cannot be sure what the expert's preferences are. In none of the examples is there a direct payment from the consumer to the expert, which means that the consumer has little ability to provide the expert with the right incentives. Furthermore, the consumer's communication of his preferences is complicated (due to the complicated nature of the issue and the consumer's ignorance that leads him to seek advice in the first place) and the expert's task is difficult. In other words, even if the expert tried to help the consumer as well as he can, there would be some likelihood of misunderstanding and error. In a static one-shot game, we should not expect useful advice in any of these situations: By the one-shot nature, an expert would optimally recommend the alternative that earns him the (highest) bonus, since the consumer has no way to punish this behavior. Knowing this, the consumer would then not even ask for advice as the recommendation would not be consistent with his preferences. However, the above examples do not usually resemble a one-shot game. Consumers repeatedly consult the same financial adviser, use the same search engine, work with the same subordinates or visit the same physician. Repeated interaction – one could call it "relationship building" – has two interesting features: First, it is well known in game theory that cooperative behavior can be sustained in repeated interactions, even if this behavior cannot be sustained in a static one-shot game. Therefore, meaningful advice might be possible because of the repeated nature of the advice situation. Second, the adviser could learn to interpret the consumer's wishes. That is, the adviser's ability to give fitting recommendations is likely to improve over time. This is because both the adviser and the consumer can observe how previous recommendations have played out, such as whether the consumer was satisfied with the product purchased (or tried to return it), whether the consumer clicked on the recommended link (and stayed on the website or subsequently purchased something there), whether the draft was pushed forward or discarded or whether the patient was cured. The success or failure of the recommendation can be used to learn how to interpret future requests from the consumer.

It should be noted that the learning we have in mind is relationship-specific. In particular, prior learning would be of little use to the consumer if he decides to switch experts. Although the consumer might also learn how to express his wishes to some extent, most of the learning seems to be on the expert's side. This paper therefore focuses on a setting where only the expert learns, and attempts to answer several questions. The most basic question is whether an equilibrium with meaningful advice is possible. The answer, unsurprisingly, is yes. The expert will give partially useful advice in equilibrium because the consumer threatens to end the relationship (and therefore the expert's opportunity to collect bonuses) if he receives bad advice for a number of periods. The key question is whether the consumer will benefit from the expert's learning. This is unclear because the consumer's outside option is not affected by the expert's learning, i.e. the expert could counteract his improved ability to give the right recommendation by recommending the product for which he receives a bonus more often. It is shown that – under certain conditions – the consumer in a certain class of simple equilibria nevertheless benefits. The reason for this is a *value effect*. The more the expert learns about the consumer, the more valuable the consumer is to the expert in the sense that the expected discounted bonus stream from that consumer is higher. The reason is that bad advice due to misunderstandings, i.e. the expert trying to give fitting advice but failing due to misunderstanding the consumer's request, can be avoided. Given the higher value, the expert will lose more if the consumer ends the relationship and is therefore generally more inclined to give good advice to avoid exactly that. This leads to a testable prediction: The probability that a relationship will end now given that it has not already ended is lower the longer the relationship lasts.[1]

The result that consumers benefit from expert learning provides a natural explanation for a puzzle that has emerged in the literature on privacy. People do not take even simple measures to anonymize their online activities. For example, most users

[1]More precisely, the probability that the relationship will end this period is lower than it was m periods ago, where $m \in \mathbb{N}$ is a number defined by the consumer's equilibrium strategy.

use a search engine like Google directly, rather than using an anonymized service that redirects their search queries through another server before forwarding them to Google (and thus anonymizing them).[2] Privacy advocates emphasize that the more information the search engine has about a user, the greater the potential for exploitation (a simple exploitation method would be to display more sponsored links). The model shows that this is not the only effect. Due to the value effect, consumers also benefit from the search engine's learning. Staying anonymous can lead to lower consumer surplus in the model of this paper. This also explains why consumers might prefer to get advice from the same person, such as having the same financial adviser at their bank whenever they go there, or staying with the same physician instead of switching every time they fall ill.

The rest of this paper is organized as follows: Section 2 discusses related literature. Section 3 presents the model and the equilibrium analysis is performed in Section 4. Section 5 deals with welfare and anonymization. Most proofs of our theoretical results can be found in Appendix A. To compare our theoretical findings with real-world behavior, we conducted a laboratory experiment. The key results from this experiment are reported in Section 6. Section 7 discusses the results of this paper, Section 8 concludes.

2. Related literature

The consumer-expert relationship we study can be reinterpreted as a relationship between a principal and a noisily informed agent. In this sense, our work is naturally related to the cheap talk literature started by Crawford and Sobel (1982) and surveyed in Sobel (2013) and Blume et al. (2020). The fact that repeated interaction can be beneficial despite the lack of commitment is reminiscent of the literature on relational contracting started by Bengt Holmstrom (Holmström, 1978, 1982). There are two notable differences. First, most of the cheap talk literature is either static or deals with reputation concerns (Sobel, 1985; Benabou and Laroque, 1992; Park, 2005). Reputation issues are not addressed in the context of this paper but are addressed in Schottmüller (2019), where a similar model is used, which, however, does not allow for learning by the expert. Second, and more importantly, the cheap talk literature deals with a different misalignment of preferences. Typically, there is a one-dimensional decision and the expert is biased in one direction, e.g. he prefers slightly higher decisions than the decision maker. The structure here is different because the expert simply has a preferred option that is independent of the consumer's optimal option. One implication of this structure is that no meaningful advice is possible in a static setting, whereas this is obviously not

[2]There are many easy-to-use services of this type, such as `https://www.startpage.com` or `https://www.privatesearch.io`.

the case in the cheap talk literature.

Li et al. (2017) analyze a repeated games setting in which the expert's and the principal's preferred projects are always distinct but the principal's project does not always exist. Moreover, there is always a default option that yields zero for all, and a disastrous project that yields $-\infty$ for all. Only the expert observes the identity of the projects, can communicate them and they are implemented if both expert and principal put effort into the same project. Our paper differs in two ways: First, both the expert's and the consumer's preferred option always exist and they can be equal. Second, the expert is not perfectly informed about the consumer's preferred option, but he receives a signal whose quality may increase over time.

The setting in Lipnowski and Ramos (2020) is probably closest to ours, since there the principal decides in each period whether to freeze the projects or delegate the project decision to the expert. The expert observes the quality of the project (high or low) and then decides whether to implement it or nothing, but the principal never learns the quality of the project. They study an intertemporal delegation rule to create incentives for the agent/expert and find that the agent represents the principal's interests only if dynamic incentives are provided. Our setting differs as (i) the agent has only noisy information and (ii) the principal does not "pause" the expert but fires him when he is dissatisfied with his advice. Furthermore, we focus on welfare dynamics in a class of simple equilibria.

Another related strand of literature is that on consumer protection in financial advice (Inderst and Ottaviani, 2012a,b, 2009). In these papers, the financial adviser is not only concerned with getting his bonus but also with the suitability of his advice. They focus on policy interventions that provide the adviser with the right incentives or payment schemes depending on whether consumers know the adviser is biased or not. In our framework, the expert is exclusively paid by his bonus and only cares indirectly about the suitability of his advice as the consumer threatens to leave him after receiving bad advice. Moreover, we model the improvement of the signal technology over time, while Inderst and Ottaviani mostly assume an exogenous and static signal.

An important application of our paper is search engines. Previous work on this market has focused mainly on ad pricing and auctioning (Edelman et al., 2007; Edelman and Schwarz, 2010; Eliaz and Spiegler, 2011) while we focus on the strategic interaction of search engine and user. More closely related is the literature on privacy in the context of search engines. Computer science has provided ways to enable fully anonymous search through encryption even when the provider has no commitment power, see Byers et al. (2004) and Çetin et al. (2016). However, results on the benefits of personalization in internet search are ambiguous. On the one hand, already Spiekermann et al. (2001) argue that people value privacy protection but are not able to take the necessary means to meet this privacy protection goal. In the same vein, Acquisti et al. (2015) have demonstrated that people are unsure how to protect their data and what parts of their data are used for what purpose. They conclude that privacy protection should be regulated because naïve people will be harmed otherwise. We add to this literature by showing that even in the absence of naïveté it is unclear whether a user should allow personalization or not. In fact, users benefit from personalization in a certain class of simple equilibria. Experimental evidence shows that users value privacy to some extent (Tsai et al., 2011; Chellappa and Sin, 2005) and that sellers can benefit more than buyers from personalization (Hillenbrand and Hippel, 2019). On the other hand, some authors have shown that providing some personal data can benefit consumers, see Xu et al. (2007); Zimmer (2008).

3. Model

The model is a dynamic game with infinite time horizon. In each period, there are two options, one of which the *consumer (C)* must choose. One of the two options fits C's needs and therefore gives him a payoff of 1 while the other option gives him a payoff of 0. C's prior is that both options are equally likely to give him a payoff of 1.

The *expert (E)* receives a private and noisy signal about which option fits C's needs. More precisely, E's signal leads to a posterior in which one option has probability $p^k > 1/2$ to fit C's needs and the other option has probability $1 - p^k < 1/2$ to fit C's needs. Without loss of generality we call the option that is more likely to fit C's needs option 1. The precision of E's signal, p^k, is an element of a finite set $P = \{p^1, p^2, \ldots, p^n\}$ with $1/2 < p^1 < p^2 < \cdots < p^n < 1$. As E learns about C's needs over time, precision improves in the following way: Whenever E recommends the option fitting C's needs, precision improves from p^k to p^{k+1} (unless $p^k = p^n$ in which case precision remains unchanged).[3]

The *expert's payoffs* are as follows: In every period, E has a *bonus option.* That is, E receives a bonus of 1 if he recommends this option to C while he receives a payoff of 0 otherwise. Each option has ex ante the same probability of being the bonus option and the identity of the bonus option is private information of E.[4]

The *timing* is as follows. In each period, E privately observes his signal and the identity of his bonus option. Then E recommends an option to C. C follows this recommendation and period payoffs realize. Both players observe whether the recommenda-

[3] The finiteness of P simplifies the exposition, but does not affect the results. As p^i cannot increase above 1, learning must eventually flatten out in the sense that precision has to converge to an upper bound as i becomes large. Finiteness of P relieves us of the notationally burdensome task of taking limits in certain proofs and allows us to use backward induction right away.

[4] Note that more options for the expert would only make the analysis more tedious without really adding anything to the model, since the expert will only decide between his bonus option and the option he considers most likely to be the fitting option for the consumer.

tion fits C's needs or not. Then, C decides whether to *end* or *continue* the game. If C ends the game, C receives an outside option V_O in the following period while E receives no payoffs in all future periods. If C continues, another period of the same game begins. Both players discount future payoffs with discount factor $\delta \in (0,1)$. Needs and bonus option are assumed to be independent of each other and across periods.

In what follows, the word *hit* (*miss*) is used to denote the event that the recommendation fits (does not fit) the consumer's needs in a given period.

To make the problem interesting, C's outside option should be neither too attractive nor too unattractive. For example, V_O should be lower than the value the consumer would receive if he had a signal of precision p^n. If this was not satisfied, C would have the dominant strategy to end the relationship immediately. The outside option should also not be too low. More precisely, we assume that V_O is higher than the value C gets when E recommends his bonus option in each period. If this did not hold, there would be a unique perfect Bayesian equilibrium in which C always continues and E always recommends his bonus option. These two conditions are stated as

$$\frac{1/2}{1-\delta} < V_O < \frac{p^n}{1-\delta}. \tag{1}$$

Before turning to the players' strategies, let us discuss some modeling choices. We assume that the recommendation itself is payoff-relevant, i.e. E receives his bonus if he recommends the bonus option and C receives his payoff if the recommendation fits his needs. Put differently, there is no real decision by C whether or not to follow the recommendation. This is not unreasonable because C has uniform beliefs and therefore cannot draw any inference from the recommendation itself about the likelihood that the recommendation fits his needs. Given that C has continued in the previous period and thereby asked for more advice, it seems logical to follow that advice. That is, there is no reason in the model to first ask for advice and then not follow it. It is also in line with certain applications, e.g. a consumer using a search engine will typically not refuse to click on a recommended link and most patients, as long as they can afford it, will take the prescribed medication. It is assumed that at the end of a period both C and E observe whether the given recommendation fitted the consumer's needs. In the examples mentioned earlier, this last assumption is reasonable: A salesperson will observe whether the consumer tries to return the product, the civil servant will observe whether his draft is pushed forward and the doctor will find out whether the patient recovers. In the search engine example, the search engine observes whether the link was clicked and – in the case of Google – to the extent that the target website uses GoogleAnalytics, csi.gstatic, GoogleAdSense or a GooglePlus button, Google also receives information about the user's subsequent behavior on the target website.

Note that the model assumes independence at several points. First, the bonus option is independent of the consumer's needs. This is one of the main differences to the cheap talk literature and appears naturally in the examples of the introduction. Second, there is some temporal independence in the sense that the consumer's needs and the bonus options are drawn independently in each period. One way to interpret this is that the requests of the consumer are unrelated, e.g. searching for an Italian restaurant in one period and for news in another period in the search engine example or suffering from different diseases in the patient-doctor example. In the financial advice example, the market environment and the set of available products may change from period to period.

As argued before, E gets to know the consumer better, so the precision of E's signal should increase over time. Depending on the application, the precision might increase either after each interaction or after each hit or not at all. It seems realistic that a fitting recommendation tells more about a consumer's preferences than a non-fitting one. The assumption made here is that the precision increases with the number of past hits and that this increase is deterministic and commonly known by C and E. That is, no learning happens after misses. The special case of no learning at all will be analyzed later as a starting point.

It is worth noting that no meaningful advice would be possible if the game was not infinitely repeated. Let us consider the static case. E has no incentive to recommend anything other than his bonus option. C therefore receives no information about which option is more likely to fit his needs. A similar situation emerges in a finitely repeated game. The static analysis applies to the last period. Since there is no meaningful communication in the last period, C should end the game after the penultimate period (regardless of history). Anticipating this, E will optimally recommend his bonus option in the penultimate period, regardless of what his signal is. Iterating this reasoning the game unravels and no meaningful advice is possible in any period. In the infinitely repeated game, the situation changes because future bonuses may motivate E to give truthful advice even if his bonus option is option 2. As there is no last period, there is no period in which these dynamic incentives break down.

What are the *strategies* of the players in this game? We assume that the players base their decision only on observed, payoff-relevant information. That is, C's decision depends only on the sequence of hits and misses in the previous periods.[5] E has to decide in each period which option to recommend. His decision depends on his posterior belief, his bonus action and the history of hits and misses. In principle, his decision could also depend on the history of bonus options, but this possibility is neglected because his current and future payoffs do not depend on this information (neither

[5]In principle, C observes the specific recommendations but since the option labels are not observed by him, he is unable to condition his strategy on these labels.

directly nor indirectly as C's strategy cannot condition on this information, which C has not observed).

In the following, we employ two commonly used equilibrium notions and compare their outcomes. Both put further restrictions on strategies. First is the *Markov equilibrium*, where strategies condition only on the actions and information of the current period and a payoff-relevant state variable. The state variable is the current precision p^k. Consequently, E's strategy is a function $s_E\colon P \times \{1, 2\} \to [0, 1]$ that assigns a probability of recommending option 1 to every $p^k \in P$ and the identity of the bonus option. C's strategy is a function $s_C\colon P \times \{hit, miss\} \to [0, 1]$ that assigns a probability of continuing the game to every $p^k \in P$ and the success of this period's recommendation.

The second notion of equilibrium is (an extension of) *grim trigger*. C continues as long as the recommendations are hits. He ends the game if m consecutive recommendations are misses for some $m \in \mathbb{N}$. E plays a best response to this strategy. Of course, it remains to be shown that C's grim trigger strategy is a best response to E's best response, but this turns out to be straightforward unless V_O is too high.

4. Analysis

In the following, we study two classes of simple equilibria and demonstrate the welfare implications. In Markov equilibria, consumers do not benefit from learning. The logic is that the consumer's outside option does not improve when the expert learns and consequently the expert will not be willing to leave him a higher surplus. In a grim trigger equilibrium, we show that the consumer does benefit from the expert's learning. However, if we extend the grim trigger concept such that the consumer does not quit after the first bad advice but after, say, two consecutive bad advice, the consumer may even lose out (for some parameter values) due to the expert's learning.

4.1. Markov equilibrium

Note first that there is always a *babbling Markov equilibrium*. In this equilibrium, E will always recommend his bonus option and C will always stop the game. Clearly, these are mutually best responses given assumption (1). Therefore, the interesting question is not whether a Markov equilibrium exists but whether a Markov equilibrium with some information transmission exists. Before answering this question in general, it is useful to analyze the case without learning where the precision of E's signal remains constant. If C does not stop the game beforehand, this situation occurs after $n - 1$ hits in our model when E's signal has precision p^n.

4.1.1. Model without learning

Without learning, the state never changes and therefore a Markov strategy will only condition on this period's information/actions. That is, a strategy for E consists of two probabilities of recommending option 1 if (i) it is the bonus option and (ii) it is not. Similarly, a strategy of C consists of two probabilities of continuing: one in case of a hit and one in case of a miss.

In equilibrium, the probability of continuing is (weakly) higher in case of a hit than in case of a miss. Otherwise, E would have an incentive to give worst possible advice, i.e. to always recommend option 2 if it is the bonus option (and possibly even if it is not) which, according to (1), automatically implies that C is better off ending the game.

Since the probability of continuing the game is higher in case of a hit than in case of a miss, it is optimal for E to recommend option 1 if option 1 is the bonus option. In this case the incentives of C and E are aligned. E's strategy can therefore be reduced to a probability α of recommending option 1 when option 2 is the bonus option.

While other equilibria can exist, we will focus on the case where C continues with probability 1 in case of a hit. Note that this provides the greatest incentive for E to be truthful. The restriction is not problematic: It is not hard to show that whenever a non-babbling Markov equilibrium exists, there exists a Markov equilibrium in which C continues with probability 1 in case of a hit. Furthermore, this is the equilibrium that Pareto dominates all other Markov equilibria. Under this constraint, C's strategy is simply a probability β of continuing in case the recommendation is a miss.

Denote E's equilibrium value, i.e. his discounted expected payoff stream at the start of a period (even before knowing the identity of the bonus option), by Π. If option 2 is the bonus option, E prefers recommending option 1 if

$$\begin{aligned} p\delta\Pi + (1-p)\beta\delta\Pi &\geq 1 + p\beta\delta\Pi + (1-p)\delta\Pi \\ \Leftrightarrow \beta &\leq \frac{(2p-1)\delta\Pi - 1}{(2p-1)\delta\Pi}. \end{aligned} \tag{2}$$

Denote C's equilibrium value by V and note that C is willing to continue only if $V \geq V_O$. Since this is independent of whether the current period's recommendation was a hit or a miss and since C continues for sure after a hit, C must either continue with probability 1 even after a miss, $\beta = 1$, or C must be indifferent, $V = V_O$. The former cannot happen in equilibrium: (2) cannot hold for $\beta = 1$ and E would therefore always recommend his bonus option. However, by (1), C would then strictly prefer not to continue. Therefore, $V = V_O$ in equilibrium and consequently α has to be such that[6]

$$\begin{aligned} V_O &= \frac{1}{2}p + \frac{1}{2}\left(\alpha p + (1-\alpha)(1-p)\right) + \delta V_O \\ \Leftrightarrow \alpha &= \frac{2(1-\delta)V_O - 1}{2p-1}. \end{aligned} \tag{3}$$

By (1), $\alpha \in (0,1)$. Hence, in an informative Markov equilibrium, E uses a mixed strategy and E is only willing to mix if (2) holds with equality. Given these equilibrium strategies one can determine the equilibrium values and obtain conditions for the existence of a non-babbling Markov equilibrium.

Proposition 1. *A non-babbling Markov equilibrium in the model without learning exists if and only if*

$$\frac{1-\delta}{\delta} \leq \frac{4p-3}{2}. \tag{4}$$

In such an equilibrium $V = V_O$ and $\Pi > 0$ and in the Pareto optimal Markov equilibrium α is given by (3) and $\beta = 1 - 1/[(2p-1)\delta\Pi]$.

Note that condition (4) is more likely to be satisfied the higher p and δ are. Moreover, it implies $p \geq 0.75$, so the signal quality has to be quite high in order to guarantee the existence of a Markov equilibrium. Intuitively, this makes sense since the expert has to be incentivized to recommend option 1 in some cases even when it is not his bonus option. This will happen when the expert is more patient (high δ) or is reasonably sure to produce a hit (high p) in this case, such that the next period will be reached with higher probability.

4.1.2. Model with learning

Also in the model with learning, it is straightforward to see that E will always recommend option 1 when option 1 is the bonus option. As before, we will focus on non-babbling Markov equilibria in which C continues for sure in case of a hit. Strategies are therefore given by sets of probabilities $\{\alpha^k\}_{k\in\{1,\dots,n\}}$ and $\{\beta^k\}_{k\in\{1,\dots,n\}}$. The players' values, i.e. their expected discounted payoff streams at the start of a period with precision p^k, are denoted by Π^k and V^k. It follows from the previous subsection that such an equilibrium can only exist if (4) holds (for $p = p^n$). This condition is necessary but not sufficient for the existence of a non-babbling Markov equilibrium and is therefore generalized below.

The first step is to show that in no period E will recommend option 1 regardless of the identity of the bonus option while C continues regardless of whether the recommendation is a hit or a miss. While this property is not surprising, it is also not

[6] As C is indifferent, we can determine his value $V = V_O$ by writing down the expected payoff stream if he continued for sure this period.

straightforward: After all, recommending option 1 gives E a higher chance to move to the next highest precision and in principle it would be possible for this to motivate him to be truthful (if Π^{k+1} is sufficiently larger than Π^k).

Lemma 1. *In Markov equilibrium $\alpha^k = \beta^k = 1$ cannot hold for any k because E's best response to $\beta^k = 1$ is $\alpha^k = 0$.*

Lemma 2. *In every Markov equilibrium $V^k = V_O$ for all $k \in \{1, 2, \ldots, n\}$.*

Lemma 2 implies E's strategy in Markov equilibrium. If the game reaches precision p^k with positive probability in a Markov equilibrium, then E has to mix such that C is indifferent between continuing and stopping. That is,

$$\begin{aligned} V_O &= \frac{1}{2}p^k + \frac{1}{2}\left(\alpha^k p^k + (1-\alpha^k)(1-p^k)\right) + \delta V_O \\ \Leftrightarrow \alpha^k &= \frac{2(1-\delta)V_O - 1}{2p^k - 1}. \end{aligned} \tag{5}$$

Note that α^k, as given by (5), is in $(0,1)$ by assumption (1). Consequently, E must be indifferent between recommending either option if the bonus option is option 2. This indifference condition determines β^k:

$$\begin{aligned} 1 + p^k\beta^k\delta\Pi^k + (1-p^k)\delta\Pi^{k+1} &= 0 + p^k\delta\Pi^{k+1} + (1-p^k)\beta^k\delta\Pi^k \\ \Leftrightarrow \beta^k &= \frac{(2p^k - 1)\delta\Pi^{k+1} - 1}{(2p^k - 1)\delta\Pi^k}. \end{aligned} \tag{6}$$

Note that Π^n is given by the stationary equilibrium value derived in the proof of Proposition 1. From this, Π^{n-1} and β^{n-1} can be obtained and by backward induction all other β^k and Π^k can also be obtained. A non-babbling Markov equilibrium exists if all such obtained β^k are in $[0,1]$. The following proposition gives a necessary and sufficient condition for exactly this.

Proposition 2. *A non-babbling Markov equilibrium in the model with learning exists if and only if*

$$\frac{\delta^{n-2}}{1-\delta}\frac{4p^n - 3}{4p^n - 2} + \sum_{k=0}^{n-3}\delta^k\frac{4p^{k+2} - 3}{4p^{k+2} - 2} \geq \frac{1}{\delta(2p^1 - 1)}. \tag{7}$$

In this Markov equilibrium, $V^k = V_O$ and

$$\Pi^k = \frac{\delta^{n-k}}{1-\delta}\frac{4p^n - 3}{4p^n - 2} + \sum_{j=0}^{n-k-1}\delta^j\frac{4p^{j+k} - 3}{4p^{j+k} - 2} \tag{8}$$

for $k \in \{1, 2, \ldots, n\}$, and α^k and β^k are given by (5) and (6), respectively.

4.2. Simple grim trigger strategies and m-equilibrium

Like most repeated games, the game described here has multiple perfect Bayesian Nash equilibria. We will now focus on a class of equilibria in which C employs the following particularly simple strategy: C continues the relationship unless the past $m \geq 1$ recommendations were misses. After m consecutive misses, C stops the game and consumes his outside option. Since this strategy is somewhat similar to the grim trigger strategies taught in introductory game theory, we will call this strategy a simple grim trigger strategy of length m or *m-strategy* for short. A perfect Bayesian Nash equilibrium in which C uses an m-strategy is called *m-equilibrium*.

When can an m-strategy be optimal for C? First, C must have a continuation value of at least V_O after any history that contains fewer than m consecutive misses. Second, continuing after m misses must result in a continuation value of at most V_O. The latter can be easily achieved: According to (1), it is optimal to end the game if E recommends his bonus option in all subsequent periods. In an m-equilibrium, continuing after m or more misses is clearly off the equilibrium path. Hence, the following off path beliefs of E will make this response optimal: If C has continued after m misses before, then E believes that C will end the game in the next period regardless of whether there is a miss or hit in the current period. Given this belief, it is clearly optimal to recommend the bonus option now. This implies that it is indeed optimal for C to end the game after m (or more) misses. These off path beliefs are not ruled out by perfect Bayesian Nash equilibrium or normal refinements.

Based on this off path construction, the following steps suffice to construct an m-equilibrium. First, derive E's best response to C's m-strategy. Second, verify that C's continuation value on the equilibrium path is at least V_O. This implies that C's strategy is optimal as ending the game earlier always yields only V_O.

What is E's best response to an m-strategy? In a given period, E is always tempted to recommend the bonus option in order to secure a payoff of 1. The downside of this choice is that a miss is quite likely if the posterior belief that the bonus action fits C's needs is low. An additional miss brings E closer to the end of the relationship, stopping the bonus stream forever and therefore leading to a payoff of zero for E. It is immediate that E will always recommend option 1 if option 1 is the bonus option.

We denote the value of the expected discounted bonus stream after t consecutive misses, when the signal strength is p^k, by Π_t^k. After $t-1$ consecutive misses, it is optimal for E to recommend option 1 instead of the bonus option (in case the two are not identical) if the following relation (9) holds.

$$\begin{aligned} p^k\delta\Pi_t^k + (1-p^k)\delta\Pi_0^{k+1} + 1 &\leq p^k\delta\Pi_0^{k+1} + (1-p^k)\delta\Pi_t^k \qquad (9) \\ \Leftrightarrow \frac{1}{\delta(2p^k-1)} &\leq \Pi_0^{k+1} - \Pi_t^k \end{aligned}$$

Note that in an m-equilibrium $\Pi_m^k = 0$. Consequently, E – for a given k – is most inclined to give good advice after $m-1$ misses. The following lemma verifies a more general result: Π_t^k is decreasing in the number of misses t which implies that E becomes more eager to give good advice as the number of misses increases. Furthermore, E benefits from learning in the sense that Π_0^k is increasing in k.

Lemma 3. *In every m-equilibrium, Π_0^k is increasing in k and Π_t^k is decreasing in t.*

Lemma 3 has a direct implication for E's strategy in an m-equilibrium: As Π_t^k is decreasing in t, (9) is more likely to be satisfied for higher t (fixing k). Thus, for a given precision p^k, E will recommend the bonus option if t is low and option 1 if t is sufficiently high (in case the two do not coincide). This result is stated as a corollary for further reference.

Corollary 1. *In every m-equilibrium, E uses a precision dependent cutoff strategy. That is, E recommends the bonus option if the number of consecutive misses t with signal strength p^k is strictly below some threshold $l^k \in \{0, 1, \ldots, m\}$ and recommends option 1 otherwise.*

Note that both the case $l^k = 0$, corresponding to E always recommending option 1, and the case $l^k = m$, corresponding to always recommending the bonus option, are allowed. For $t \geq l^k$, E's value can be written as $\Pi_t^k = 1/2 + p^k \delta \Pi_0^{k+1} + \left(1 - p^k\right) \delta \Pi_{t+1}^k$. Keeping in mind that $\Pi_m^k = 0$ in an m-equilibrium, backward induction gives for $t \in \{l^k, \ldots, m-1\}$

$$\Pi_t^k = \sum_{j=0}^{m-t-1} \delta^j \left(\frac{1}{2}(1-p^k)^j + p^k(1-p^k)^j \delta \Pi_0^{k+1} \right). \tag{10}$$

For $t < l^k$, E's value is $\Pi_t^k = 1 + \delta \Pi_0^{k+1}/2 + \delta \Pi_{t+1}^k/2$. Using the expression for $t \geq l^k$ above, iterating backwards yields for $t < l^k$

$$\Pi_t^k = \sum_{j=0}^{l^k-t-1} \delta^j \left(\left(\frac{1}{2}\right)^j + \left(\frac{1}{2}\right)^{j+1} \delta \Pi_0^{k+1} \right) + \sum_{j=0}^{m-l^k-1} \left(\frac{\delta}{2}\right)^{l^k-t} \delta^j \left(\frac{1}{2}(1-p^k)^j + p^k(1-p^k)^j \delta \Pi_0^{k+1} \right). \tag{11}$$

Using relation (10), we can derive the exact value of the threshold l^k:

Lemma 4. *The threshold l^k chosen by E in an m-equilibrium is given by*

$$l^k = \begin{cases} 0, & \text{if } \frac{1}{2} + p^k\delta\Pi_0^{k+1} \leq (1-\delta(1-p^k))(\Pi_0^{k+1} - \frac{1}{(2p^k-1)\delta}) \\ m, & \text{if } \Pi_0^{k+1} < \frac{1}{(2p^k-1)\delta} \\ \max\left\{0, \left\lceil m - 1 - \frac{\ln\left(1-(1-\delta(1-p^k))\frac{\Pi_0^{k+1}-\frac{1}{(2p^k-1)\delta}}{\frac{1}{2}+p^k\delta\Pi_0^{k+1}}\right)}{\ln\left(\delta(1-p^k)\right)}\right\rceil\right\}, & \text{else.} \end{cases} \tag{12}$$

Note that Lemma 4 also implies that $m > l^k$ always holds in the third case, since the logarithm in the numerator is negative (the negations of the first two conditions ensure that the term inside the logarithm is between 0 and 1). This justifies the following

Remark 1. *If $l^k = m$ for some k in an m-equilibrium, then also $l^i = m$ for all $i \in \{1, \dots, k-1\}$. This follows directly from Lemma 4 as Π_0^{k+1} is increasing in k and $\frac{1}{(2p^k-1)\delta}$ is decreasing in k.*

It is useful to first analyze the case without (further) learning which occurs after $n-1$ hits.

4.2.1. Model without learning

For $k \geq n$, $\Pi_0^k = \Pi_0^n$ since there is no more additional learning. This implies that in an m-equilibrium, Π_0^n has to solve (11) with the same Π_0^n on both sides of the equation. Furthermore, l^n in this equation has to be optimal in the sense of (9). The following lemma implies that there exist unique Π_0^n and l^n satisfying these optimality conditions.[7]

Lemma 5. *E has a unique best response to C's m-strategy in the model without learning.*

Whether an m-equilibrium exists depends on C's outside option. If E's best response to C's m-strategy, as derived in the proof of Lemma 5, leaves C with a sufficiently high value after 0 misses, then an m-equilibrium exists.

Proposition 3. *An m-equilibrium in the model without learning does not exist if*

$$2p^n - 1 < \frac{1-\delta+(\delta/2)^{m+1}}{\delta(1-(\delta/2)^m)}. \tag{13}$$

If (13) does not hold, an m-equilibrium exists if and only if $V_O \leq \bar{V}_O$ for some $\bar{V}_O$ satisfying (1).[8]

[7] For uniqueness, we require the tie-breaking rule that E recommends option 1 if he is indifferent. Without this tie-breaking uniqueness is (only) generic.

[8] Note that for $m = 1$, the condition (13) reduces to $2p^n - 1 < \frac{1-\frac{\delta}{2}}{\delta} \Leftrightarrow p^n < \frac{3}{4} + \frac{1-\delta}{2\delta}$. This is exactly the existence condition (4) for a Markov equilibrium without learning.

4.2.2. Model with learning

We start by deducing explicit formulas for the continuation value V_t^k of the consumer after t consecutive misses and with precision k. As $V_m^k = V_O$ in an m-equilibrium, the value for $t \in \{l^k, \dots, m-1\}$ can be derived by backward induction. We obtain

$$V_t^k = \sum_{j=0}^{m-t-1} (1-p^k)^j \delta^j p^k (1+\delta V_0^{k+1}) + (1-p^k)^{m-t}\delta^{m-t} V_O. \tag{14}$$

Using this, we can also derive the value of V_t^k for $t < l^k$. It is given by

$$\begin{aligned} V_t^k = &\sum_{j=0}^{l^k-t-1} \left(\frac{\delta}{2}\right)^j \frac{1}{2}(1+\delta V_0^{k+1}) \\ &+ \left(\frac{\delta}{2}\right)^{l^k-t} \left(\sum_{j=0}^{m-l^k-1} (1-p^k)^j \delta^j p^k (1+\delta V_0^{k+1}) + (1-p^k)^{m-l^k} \delta^{m-l^k} V_O \right). \end{aligned} \tag{15}$$

Before we compute the expert's expected value Π_0 at the start of the game, we introduce some notation. In an advice relationship between a consumer and an expert, let $w = (w_1, \dots, w_{n-1})$ denote the vector of waiting times until the first, second, …, $(n-1)-$th hit, where w_i denotes the number of periods in learning level i (with precision p^i) until the $i-$th hit occured. In an m-equilibrium, $w_i > m$ implies that the consumer will fire the expert as he produced at least m consecutive misses. Hence, there are two types of possible histories in an advice relationship: First, those where the expert produced at least $n-1$ hits and reached the last precision level p^n. Second, those where the expert produced at least m consecutive misses before p^n was reached. We denote these two sets of histories by

$$\mathcal{W}_n = \{w = (w_1, \dots, w_{n-1}) \in \mathbb{N}^{n-1} | 1 \le w_i \le m \ \forall i \in \{1, \dots, n-1\}\} \text{ and}$$
$$\mathcal{W}_f = \{w = (w_1, \dots, w_{j^*}) \in \mathbb{N}^{j^*} \text{for some } 1 \le j^* \le n-1 | w_{j^*} = m+1, 1 \le w_i \le m \ \forall i < j^*\}.$$

The set of all feasible histories in an m-equilibrium is then given by $\mathcal{W} = \mathcal{W}_n \cup \mathcal{W}_f$. For any $w \in \mathcal{W}_f$, let us denote by $len(w)$ the dimension of the vector w. This value always corresponds to the learning level in which the expert gets fired because he produces m consecutive misses. We can now derive Π_0.

Proposition 4. *In an m-equilibrium, let $(l^k)_{k=1,\dots,n}$ denote the vector of switching strategies for the expert, depending on the precision level. The expected value Π_0 of the expert* *at the beginning of the game is given by the formula*

$$\Pi_0 = \sum_{\bar{w}\in\mathcal{W}} \mathbb{P}(w=\bar{w})\mathbb{E}(\Pi_0|w=\bar{w}),$$

$$where\ \mathbb{P}(w=\bar{w}) = \begin{cases} \prod_{i=1}^{n-1}\left(\mathbb{1}_{\{\bar{w}_i\le l^i\}}(\frac{1}{2})^{\bar{w}_i} + \mathbb{1}_{\{\bar{w}_i>l^i\}}(\frac{1}{2})^{l^i}(1-p^i)^{\bar{w}_i-l^i-1}p^i\right), & if\ \bar{w}\in\mathcal{W}_n \\ (\frac{1}{2})^{l^{len(\bar{w})}}(1-p^{len(\bar{w})})^{m-l^{len(\bar{w})}} * & \\ \prod_{i=1}^{len(\bar{w})-1}\left(\mathbb{1}_{\{\bar{w}_i\le l^i\}}(\frac{1}{2})^{\bar{w}_i} + \mathbb{1}_{\{\bar{w}_i>l^i\}}(\frac{1}{2})^{l^i}(1-p^i)^{\bar{w}_i-l^i-1}p^i\right), & if\ \bar{w}\in\mathcal{W}_f \end{cases}$$

$$and\ \mathbb{E}(\Pi_0|w=\bar{w}) = \begin{cases} \sum_{k=1}^{n-1}\left(\mathbb{1}_{\{\bar{w}_k\le l^k\}}\sum_{h=0}^{\bar{w}_k-1}\delta^h + \mathbb{1}_{\{\bar{w}_k>l^k\}}\left(\sum_{h=0}^{l^k-1}\delta^h + \frac{1}{2}\sum_{h=l^k}^{\bar{w}_k-1}\delta^h\right)\right)\delta^{\sum_{j=1}^{k-1}\bar{w}_j} & \\ +\delta^{\sum_{j=1}^{n-1}\bar{w}_j}\frac{\sum_{g=0}^{l^n-1}(\frac{\delta}{2})^g+\sum_{g=l^n}^{m-1}\delta^g(\frac{1}{2})^{g-l^n}}{1-\sum_{g=0}^{l^n-1}(\frac{\delta}{2})^{g+1}-\sum_{g=l^n}^{m-1}(\frac{1}{2})^{l^n}\delta^g(1-p^n)^{g-l^n}p^n}, & if\ \bar{w}\in\mathcal{W}_n \\ \sum_{k=1}^{len(\bar{w})-1}\left(\mathbb{1}_{\{\bar{w}_k\le l^k\}}\sum_{h=0}^{\bar{w}_k-1}\delta^h + \mathbb{1}_{\{\bar{w}_k>l^k\}}\left(\sum_{h=0}^{l^k-1}\delta^h + \frac{1}{2}\sum_{h=l^k}^{\bar{w}_k-1}\delta^h\right)\right)\delta^{\sum_{j=1}^{k-1}\bar{w}_j} & \\ +\delta^{\sum_{j=1}^{len(\bar{w})-1}\bar{w}_j}\left(\sum_{h=0}^{l^{len(\bar{w})}-1}\delta^h + \frac{1}{2}\sum_{h=l^{len(\bar{w})}}^{m-1}\delta^h\right), & if\ \bar{w}\in\mathcal{W}_f. \end{cases}$$

The following result deals with the hazard rate, i.e. the probability that the expert is fired in a given learning level, conditional on having reached that level. More concretely, we denote by $HR(k)$ the probability of having m consecutive misses in an m-equilibrium after reaching precision level k. For the last precision level $k=n$, $HR(n)$ denotes the probability of being fired in this level without having scored a hit before (since the game is infinitely repeated, the probability of being fired in the last precision level is 1).

Proposition 5. *If* $p^{k+1} \ge 1-(1-p^k)^m 2^{m-1}$ *holds for all* k *in* $\{1,\dots,n-1\}$, *then* $HR(k)$ *is decreasing in* k *in an* m*-equilibrium.*

Example 1. *To illustrate the above proposition, let us consider an* $m=2$ *equilibrium with an initial precision of* $p^1=0.51$. *The subsequent precision levels that guarantee a decreasing hazard rate are given by* $p^2=0.5198, p^3\approx 0.5388, p^4\approx 0.5746, p^5\approx 0.6381, p^6\approx 0.7380, p^7\approx 0.8627, p^8\approx 0.9623, p^9\approx 0.9972.$

Example 2. *Figure 1 shows the precision levels that ensure a decreasing hazard rate according to Proposition 5 for* $m=2, m=3$ *and* $m=4$.

5. Welfare dynamics and anonymization

In this section, we discuss the dynamics of consumer surplus. Since the consumer's value equals his outside option regardless of the precision level in Markov equilibrium, the consumer does not benefit from learning in a Markov equilibrium. This is consistent with the argument that the expert can pocket all the benefit since the consumer's outside option (and therefore bargaining position) does not improve as the expert learns. However, analysis of m-equilibria shows that this logic may be flawed. Consider first

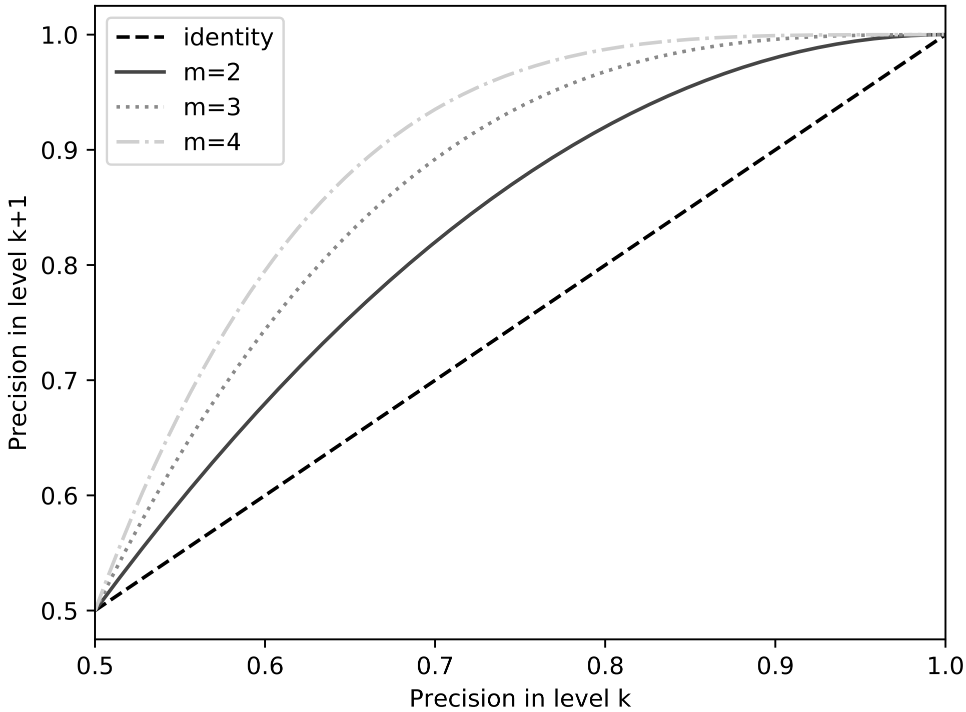


Figure 1: Sufficient learning jumps for a decreasing hazard rate

the case of a classic grim trigger strategy, i.e. $m = 1$. The follwing proposition implies that consumers benefit from learning in this class of equilibria.

Proposition 6. *In an $m = 1$ equilibrium, V_0^k is strictly increasing in k and l^k is weakly decreasing in k.*

That is, the consumer can benefit for two reasons: simply because the expert's precision and therefore the advice quality improves but also because the expert's strategy can become more favorable over time. The intuition is that, by Lemma 3, the expert's profits are increasing in the precision level k (as long as the game continues). Therefore, as precision increases, he is more inclined to give good advice in order to reap the increasing future benefits.

To further illustrate the previous result and also to shed light on the dynamics in m-equilibria for $m > 1$, we now consider the case of only two precision levels, p^1 and p^2. We are primarily interested in the expert's choice of optimal thresholds l^1 and l^2, since they determine the distribution of welfare between the consumer and the expert. First, we study the $1-$equilibrium in which the consumer ends the relationship after the first miss. The following lemma shows that the expert's choice depends on how large δ is relative to p^1 and p^2.

Lemma 6. *Let $n = 2$. If the consumer ends the game after one miss, the expert's advice choices l^1, l^2 are given by*

$$(l^1, l^2) = \begin{cases} (1,1), & \textit{for } \delta < \frac{1}{2p^2 - 1/2} \\ (1,0), & \textit{for } \frac{1}{2p^2-1/2} < \delta < \frac{1}{p^1+p^2-1/2} \\ (0,0), & \textit{for } \delta > \frac{1}{p^1+p^2-1/2} \end{cases} \tag{16}$$

The previous lemma implies that consumers benefit from learning in the $m = 1$ equilibrium: In equilibrium, l^2 (or more generally l^n for n rounds of learning) must equal 0. Otherwise, the consumer would be better off ending the advice relationship once the last precision level is reached, i.e. the $m = 1$ strategy would not be a best response. This implies that the cutoffs l^k are weakly decreasing in k in the $n = 2$ case and therefore advice improves in k for two reasons. First, the consumer can benefit from a lower l^k and thus a more honest advice strategy from the expert. Second, even if $l^1 = l^2 = 0$ and therefore the expert's advice strategy remains constant, the consumer benefits from learning as the signal technology improves.

While consumers benefit from learning in an $m = 1$ equilibrium, this is not necessarily the case in an $m > 1$ equilibrium. We illustrate this in the simplest possible case, i.e. only one round of learning ($n = 2$) and $m = 2$. In this case, we show that there are parameter values for which $l^1 = 0$ and $l^2 = 1$ in an $m = 2$ equilibrium. That is, the expert is less willing to give good advice after the signal technology improved. In our example, this change in expert strategy affects the consumer's payoff more than the improvement in signal technology and therefore the consumer's value will be lower at the beginning of a period with improved signal technology than at the beginning of the game.

Lemma 7. *Let $n = 2$. In an $m = 2$ equilibrium, $l^1 = 0$ and $l^2 = 1$ if and only if both*

$$\frac{1+\delta/4}{1-\delta/2-\delta^2p^2/2} > \max\left\{\frac{1+\delta/2}{1-\delta/2-\delta^2/4}, \frac{(1+\delta)/2-\delta^2p^2/2}{1-p^2\delta-(1-p^2)\delta^2p^2}\right\}$$
$$\textit{and } \frac{1}{2} + \frac{(1-p^1)\delta}{2} + \frac{p^1\delta(1+(1-p^1)\delta)(1+\delta/4)}{1-\delta/2-\delta^2p^2/2}$$
$$\geq \max\left\{1+\delta/4+\frac{(\delta/2+\delta^2p^1/2)(1+\delta/4)}{1-\delta/2-\delta^2p^2/2}, 1+\delta/2+\frac{(\delta/2+\delta^2/4)(1+\delta/4)}{1-\delta/2-\delta^2p^2/2}\right\} \textit{ hold.}$$

Thus, an $m = 2$ equilibrium with $l^1 = 0$ and $l^2 = 1$ exists if the two inequalities above hold simultanenously. For $\delta = .98$, $p^1 = .85$ and $p^2 = .95$ both inequalities hold with strict inequality. Furthermore, for $V_O = 30$, we get $V_0^2 \approx 31.33 < 31.56 \approx V_0^1$, which proves the following result:

Proposition 7. *There exists an open set of parameters such that $V_0^1 > V_0^2$ in an $m = 2$ equilibrium.*

What is the intuition behind these results? Let us first consider Proposition 6. Knowing the consumer better means that the expert is better able to keep the consumer satisfied. The expert's value from giving good advice is higher when the signal is better because he is less likely to lose the consumer due to random errors. However, the value of recommending the bonus option does not depend on the signal technology. Thus, improvements in signal technology make it relatively more attractive to give good advice. Technically, the better the signal technology, the higher the continuation value of the relationship for the expert. This means that future payoffs and a continuation of the relationship gain in importance when the expert's signal technology improves and he is therefore more willing to give good advice. We call this the *value effect* of improved information and note that this effect is positive for the consumer.

Proposition 7 illustrates another dynamic effect that comes into play in more complicated equilibria. If the expert does not expect the consumer to end the advice relationship in case of a miss, it may be optimal for the expert to gamble: recommend the bonus option today and hope, in case of a miss, that recommending option 1 tomorrow will prevent the consumer from ending the relationship. The better the expert gets to know the consumer, the greater the incentive to gamble: The improved signal means that it is more likely that he will be able to provide a good recommendation if that is what is needed to keep the consumer tomorrow. Put differently, the risk of ending the relationship is lower because the expert can be reasonably confident of providing a fitting recommendation "on the spot" if this is needed to keep the consumer. This *gambling effect* is negative for the consumer. In the example above, the gambling effect outweighs the value effect, so the consumer's continuation value is higher when the signal technology is worse. Note that the gambling effect is not present in $m = 1$ equilibria, since in such an equilibrium the consumer ends the advice relationship immediately after the first miss.

We will now turn to the question of anonymization. The use of anonymized services makes relationship-specific learning impossible. For example, an internet search engine cannot personalize search results if the consumer uses an anonymized version of the search engine.[9] In our model, anonymization corresponds to facing an expert who always remains at the precision level p^1 due to his inability to learn. Will the consumer benefit from anonymization? In a Markov equilibrium, the consumer surplus is always equal to the outside option, so anonymization has no effect. In an $m = 1$ equilibrium, on the other hand, anonymization harms the consumer: such an equilibrium exists only

[9] Anonymized versions of major internet search engines are widely available, see for example `https://www.startpage.com`.

if E always recommends option 1 and in this case it is clear that the consumer would lose from anonymization.[10] However, the consumer can benefit from anonymization in m-equilibria with $m > 1$. Consider the example above, on which Proposition 7 was based. An equilibrium with $m = 2$ also exists in the game where no learning is possible due to anonymization. In this anonymization equilibrium, $l = 0$ and the consumer value is $V_0^1 = 36.44$ which is larger than V_0^1 and V_0^2 in the equilibrium without anonymization. The intuition is that in this example learning leads to gambling, i.e. when precision equals p^2, E is sufficiently confident that he can produce a hit on demand. Hence, he finds it optimal to recommend the bonus option in case the last recommendation was a hit. Without learning, the precision is too low to allow E to gamble and C benefits from sincere advice (albeit with a lower precision). This establishes the following result.

Proposition 8. *In Markov equilibrium anonymization neither harms nor benefits the consumer. In $m = 1$ equilibria the consumer always loses from anonymization while in $m > 1$ equilibria the consumer can benefit from anonymization for certain parameter values.*

6. Experimental Design and Results

In this section, we present the design of our laboratory experiment and its main results. Additional results and robustness checks can be found in Appendix C.

6.1. Experimental Design

The experiment was conducted between December 2021 and February 2022 at the Cologne Laboratory for Economic Research, University of Cologne. We used the experimental software oTree (Chen et al. (2016)) and recruited participants via ORSEE (Greiner (2015)). The study was preregistered in the AEA RCT Registry (Gramb and Schottmüller (2022)), its unique identifying number is: AEARCTR-0008682. Participants were randomly assigned to either the control group or the treatment group. In both groups, participants first read the instructions for their group, see Appendix B (in German), and answered a set of incentivized control questions. Then, players were randomly assigned the role of expert or consumer. Framing of roles was neutral in instructions and experiment. Subsequently, they played ten supergames (seven supergames in the pilot session in December, which was a treatment group session) of the game described in Section 3, each in their assigned role. After each supergame, each participant was randomly matched with another participant with the opposite role for the next supergame. The discounting of payoffs in the experiment was simulated by

[10] Note that the existence of an $m = 1$ equilibrium with anonymization implies that E will always recommend option 1 in the $m = 1$ equilibrium without anonymization: this follows directly from (9) and the facts that $\Pi_t^k = 0$ for $t > 0$ in $m = 1$ equilibrium and Π_0^{k+1} is weakly larger with learning than without.

an exogenous stopping probability. After each round of a supergame, the game was exogenously ended with a probability of 10%, corresponding to a value of $\delta = 0.9$ in our model. In the control group, experts had a constant signal strength of 0.82. That is, the control group can be interpreted as a setting in which advice is given anonymously and therefore learning is not possible. In the treatment group, the first precision level was also 0.82 and precision was increased by 0.02 after each hit up to a maximum precision of 0.9. That is, the treatment group represents a setting in which personalized advice and incremental learning is possible. Once a consumer decided to end the game in either group, he immediately received a payoff of 5 points (while the expert's bonus and the consumer's payoff in case of a hit were both 1 point). It should be noted that in our model the outside option is paid out at the beginning of the next period (since it always exists). Thus, the payout of 5 points after firing the expert corresponds to an outside option of $V_O = \frac{5}{\delta} = \frac{50}{9} = 5.\bar{5}$. After all supergames were completed, one supergame was randomly selected for each participant and the points earned there were paid out (with one point being worth 1€). At the end, participants were asked incentivized questions eliciting their risk attitude and completed a non-incentivized survey about trust attitude, age, gender and faculty. Additionally, each participant was paid a show-up fee of 4€. Participants' total payments ranged from 4€ to 25€. One session lasted between 29 and 56 minutes. There were seven sessions with a total of 156 participants in the treatment group and four sessions with a total of 98 participants in the control group. No participant attended more than one session.

6.2. Results

The main outcomes we are interested in are advice quality and consumer welfare in both groups. Let us start with advice quality. We measure this as the share of good advice given by the expert (in terms of the recommendation of option 1) in all situations where he faced a tradeoff (bonus option was option 2). Figure 2 shows that the advice quality in the treatment group is significantly better than in the control group. Hence, the potential increase in learning level incentivizes the experts to give better advice to retain consumers. As can be seen in Figure 3, this expert behavior leads to higher average consumer welfare in the treatment group, although the difference is not significant. A possible reason for the (only) small increase in consumer welfare in the treatment group is that consumers tend to distrust the expert more at higher learning levels. This can be seen in the firing rates in the treatment group.

In Figure 4, we see that the hazard rate[11] increases overall with learning level. Specifically, the hazard rate for precision levels p_2, p_3 and p_4 is significantly higher than for lower levels p_0 and p_1. This could also drive the effect seen in Figure 2, where

[11] In this general case, the hazard rate is simply the relative frequency with which consumers fire experts at a given learning level.

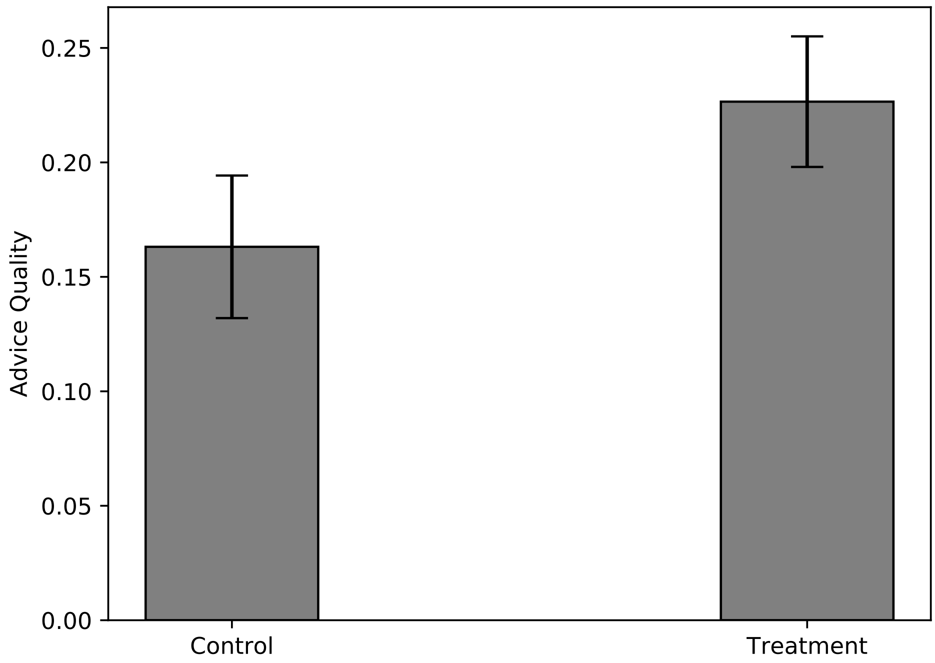


Figure 2: Advice Quality in Control and Treatment group

experts try to *convince* consumers not to fire them by giving them even more good advice. Note that this behavior does not contradict our theoretical predictions: The sufficient condition from Proposition 5 that the hazard rate decreases would require a level $p_1 \geq 0.9352$ for the value $p_0 = 0.82$. In the experiment, we set $p_1 = 0.84$, which is too small for the model to predict a decreasing hazard rate. One possible reason for this increase in hazard rate could be attribution of failure: At low learning levels, the consumer might attribute a miss to the expert's low signal strength. At high learning levels, it becomes increasingly likely that a miss is due to the expert's strategy to collect his bonus instead of giving good advice. This is then punished by the consumer who fires such experts. Interestingly, the hazard rate in the control group is significantly higher than the hazard rate for the first two learning levels in the treatment group. This suggests that consumers assume that the learning incentive has a positive effect on the relationship and do not fire the expert to establish such a long-term relationship.

7. Discussion

The results in Section 4 and 5 have implications for anonymization. Activists and experts alike recommend measures to preserve anonymity online. Although many of these recommendations are easy to follow, such as using an anonymized version of Google

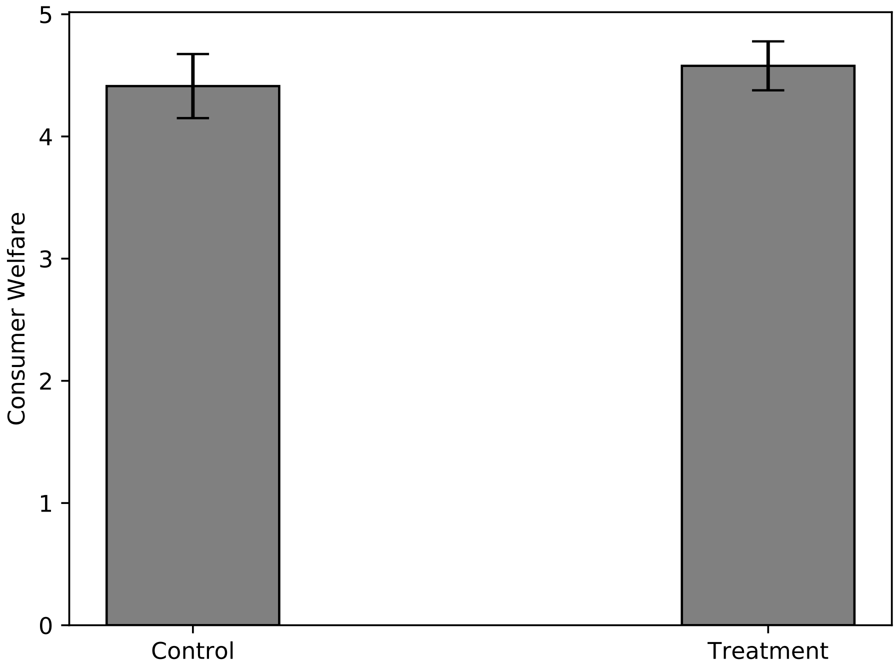


Figure 3: Consumer Welfare in Control and Treatment group

instead of Google itself, hardly any internet user follows them. The above analysis indicates that consumers might be right not to anonymize: Personalized recommendations are more valuable not only from a total surplus perspective, but also from a consumer perspective in m-equilibrium if $m = 1$ (and often also if $m > 1$). The reason for this is simple. The more past usage data is available, the more valuable the customer is. The expert, e.g. Google, does not want to risk losing valuable customers. Hence, a customer enjoys better service when the expert can use past usage data from him. This theoretical finding is also supported by our experimental results: As we have shown in Section 6, experts give better advice in the treatment group where it is possible to learn from past interactions.

The same principle applies to other applications than Google and explains why long-term advisers are more valuable than short-term advisers. The m-equilibrium provides an interesting prediction for the hazard rate, i.e. the probability that a consumer will end the relationship after a certain number of hits if he has not already ended it. In an m-equilibrium, the hazard rate decreases over time when the change in signal quality between two learning levels is sufficiently high.

Of course, these results are subject to some caveats. The first is that the outside option of the consumer was held constant. If the outside option is an alternative expert,

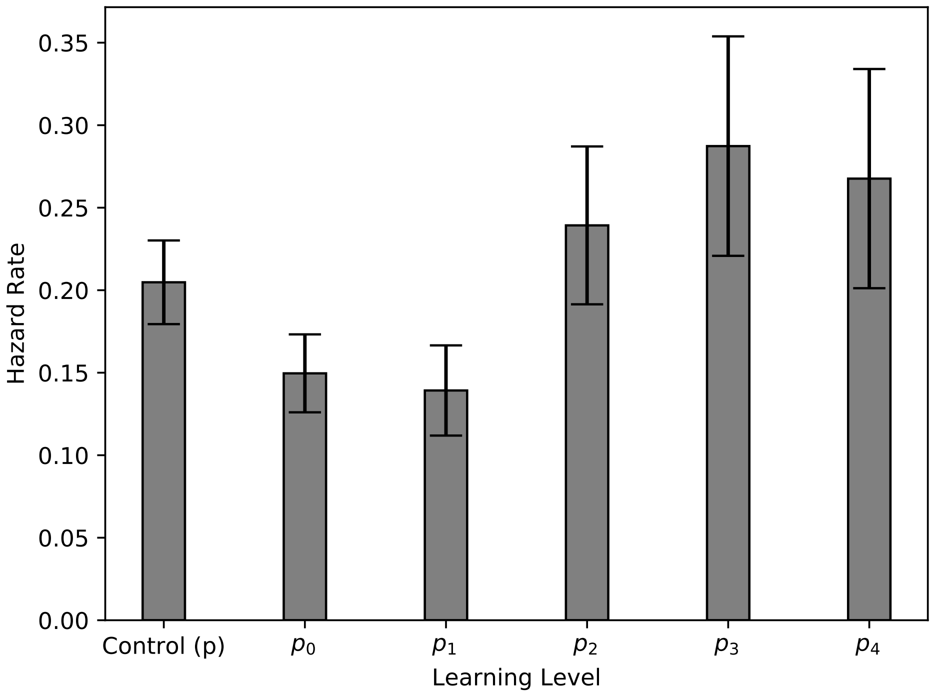


Figure 4: Hazard Rate per Learning Level in Control and Treatment group

this could change. To give an example, say there are two experts and everyone agrees that Expert 1 is slightly more knowledgeable than Expert 2. The outside option then corresponds to getting advice from Expert 2. If everyone uses Expert 1, however, Expert 2 might be out of business and take up a different job. In the long term, the outside option might therefore decline and eventually drop below $1/(2-2\delta)$. In this case, the unique equilibrium is that the expert recommends his bonus action in each period and consumers would suffer. However, an m-equilibrium is not sensitive to lower outside options as long as the outside option remains above $1/(2-2\delta)$.

Another caveat, particularly in the context of anonymizing online activity, is that the model does not address potential extortion arising from abuse of data outside the advice relationship. According to the model, a customer benefits from personalized advice and a prerequisite for such personalized advice is that data about past interactions be stored. If this data gets into the hands of a third party, it could be used by that third party against the consumer; think health or financial records. Such third-party extortion is beyond the scope of this paper.

Another interesting result was given in Proposition 7 as it showed that too much past data can also reduce the consumer's utility (although it is always higher than his outside option). Consequently, whether anonymization is optimal or not is ambiguous and depends on the particular equilibrium played as well as the model parameters.

8. Conclusion

In this paper, we have studied an expert-consumer relationship in which the expert gets to know the consumer over time and in this way can give better advice as the relationship progresses. We have shown that this learning opportunity can be beneficial to both the consumer and the expert by introducing m-equilibria as a generalization of simple grim-trigger strategies. Empirical evidence from our laboratory experiment suggests that experts do indeed give better advice when learning is possible. However, the consumer must be aware that too much learning on the part of the expert can be detrimental to consumer welfare. The choice of how much and what data to disclose is therefore a difficult one.

Appendix

A. Proofs

Proof of Proposition 1. With $\beta = 1 - 1/[(2p-1)\delta\Pi]$, it is straightforward to determine Π:[12]

$$\begin{aligned} \Pi &= \frac{1}{2} + (p + \beta(1-p))\,\delta\Pi = \frac{1}{2} + \left(p + (1-p) - \frac{1-p}{(2p-1)\delta\Pi}\right)\delta\Pi \\ \Leftrightarrow \Pi &= \frac{4p-3}{(1-\delta)(4p-2)}. \end{aligned}$$

Plugging this back into (2) (with equality) yields

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

A non-babbling Markov equilibrium exists if $\beta^* \in [0,1]$ which is the case if and only if

$$\frac{1-\delta}{\delta} \leq \frac{4p-3}{2}.$$

□

Proof of Lemma 1. Suppose $\beta^k = 1$ and distinguish between the two cases of either option 1 or option 2 being E's bonus option (both happen with probability $\frac{1}{2}$). E's value as a function of α is then

$$\begin{aligned} \Pi^k &= \frac{1}{2}\left(p^k\delta\Pi^{k+1} + (1-p^k)\delta\Pi^k\right)(1+\alpha) + \frac{1}{2} + \frac{1}{2}(1-\alpha)\left(p^k\delta\Pi^k + (1-p^k)\delta\Pi^{k+1} + 1\right) \\ &= \frac{1}{2}(2-\alpha) + \frac{1}{2}\delta\Pi^{k+1}(2p^k\alpha + 1 - \alpha) + \frac{1}{2}\delta\Pi^k(1 + \alpha - 2p^k\alpha) \\ \Leftrightarrow \Pi^k &= \frac{2-\alpha}{2-\delta-\delta\alpha+2p^k\delta\alpha} + \frac{\delta-\delta\alpha+2p^k\delta\alpha}{2-\delta-\delta\alpha+2p^k\delta\alpha}\Pi^{k+1}. \end{aligned}$$

This implies

$$\begin{aligned} \Pi^k_{\alpha=0} &= \frac{2}{2-\delta} + \frac{\delta}{2-\delta}\Pi^{k+1} \\ \Pi^k_{\alpha=1} &= \frac{1}{2(1-\delta+p^k\delta)} + \frac{2p^k\delta}{2(1-\delta+p^k\delta)}\Pi^{k+1} \end{aligned}$$

where $\Pi^k_{\alpha=1}$ is E's equilibrium value in the supposed equilibrium (where E uses the strategy $\alpha^k = 1$) and $\Pi^k_{\alpha=0}$ is a deviation value that E would obtain if he deviated from the supposed equilibrium strategy by choosing $\alpha^k = 0$ (without changing his strategy

[12] As E is indifferent between recommending option 1 and recommending option 2 in case option 2 is his bonus option, his value is as if he always recommended option 1.

for $k' \neq k$). For $\alpha^k = 1$ to be optimal $\Pi^k_{\alpha=1} \geq \Pi^k_{\alpha=0}$ has to hold. However, it is now shown that $\Pi^k_{\alpha=0} > \Pi^k_\alpha$ for any $\alpha > 0$. This inequality can be written as

$$\frac{2}{2-\delta} + \frac{\delta}{2-\delta}\Pi^{k+1} > \frac{2-\alpha}{2-\delta-\delta\alpha+2p^k\delta\alpha} + \frac{\delta-\delta\alpha+2p^k\delta\alpha}{2-\delta-\delta\alpha+2p^k\delta\alpha}\Pi^{k+1}$$
$$\Leftrightarrow 4-2\delta-2\delta\alpha+4p^k\alpha\delta+(2-\delta-\delta\alpha+2p^k\alpha\delta)\delta\Pi^{k+1}$$
$$> 4+\alpha\delta-2\delta-2\alpha+(2+\alpha\delta-\delta-2\alpha+4p^k\alpha-2p^k\delta\alpha)\delta\Pi^{k+1}$$
$$\Leftrightarrow -3\alpha\delta+2\alpha+4\alpha p^k\delta > (1-\delta)(4p^k\alpha-2\alpha)\delta\Pi^{k+1}.$$

The latter inequality is true for all $\alpha > 0$ because Π^{k+1} is bounded from above by $1/(1-\delta)$ (which would be E's discounted payoff stream if he always recommended his bonus option and C always continued) and the previous inequality holds with $1/(1-\delta)$ in place of Π^{k+1}:

$$\begin{aligned} -3\alpha\delta+2\alpha+4\alpha p^k\delta &> (1-\delta)(4p^k\alpha-2\alpha)\frac{\delta}{1-\delta} \\ \Leftrightarrow \alpha(2-\delta) &> 0. \end{aligned}$$

This shows that $\alpha^k = 0$ is the only best response to $\beta^k = 1$ and therefore $\Pi^k_{\alpha=1} < \Pi^k_{\alpha=0}$. Consequently, $\beta^k = \alpha^k = 1$ cannot be an equilibrium. □

Proof of Lemma 2. Proposition 1 implies $V^n = V_O$. Suppose $V^k > V_O$ for some k and let k' be the highest such k. Then $\alpha^{k'}$ must be sufficiently high in order to yield a higher expected payoff than $(1-\delta)V_O$ to C in every period with precision $p^{k'}$. Now consider C's decision problem after a miss in a period with precision $p^{k'}$. As $V^{k'} > V_O$ by the definition of k', C strictly prefers to continue. Hence, $\beta^{k'} = 1$. However, E's best response to $\beta^{k'} = 1$ is $\alpha^{k'} = 0$, see the proof of Lemma 1. But given that $V^k = V_O$ for all $k > k'$ by the definition of k' and given that $\alpha^{k'} = 0$ clearly $V^{k'} < V_O$ contradicting the definition of k'. Hence, $V^k > V_O$ cannot happen for any k in equilibrium. As C can always guarantee himself a payoff of V_O by ending the game, this concludes the proof. □

Proof of Proposition 2. As E is mixing in a non-babbling Markov equilibrium when the bonus option is option 2, his value will equal the value he would get if he always recommended option 1 (keeping C's strategy fixed):

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

Plugging (6) in for β^k yields

$$\begin{aligned}\Pi^k &= \frac{1}{2} + p^k\delta\Pi^{k+1} + (1-p^k)\delta\Pi^{k+1} - \frac{1-p^k}{2p^k-1} \\ \Leftrightarrow \Pi^k &= \delta\Pi^{k+1} + \frac{4p^k-3}{4p^k-2}.\end{aligned}$$

Recall from the proof of Proposition 1 that $\Pi^n = (4p^n-3)/[(4p^n-2)(1-\delta)]$. Using this as a starting point for backward induction in the previous equation yields (8).

Next we will show that Π^k is strictly increasing in k. Let $h(p^k) = (4p^k-3)/(4p^k-2)$ and note that $h' > 0$ for $p^k \in (1/2, 1]$. To start, we show by induction that $(1-\delta)\Pi^k \geq h(p^k)$. This is obviously true for $k = n$. Now suppose $(1-\delta)\Pi^k \geq h(p^k)$ is true for all $k \geq j+1$, then $(1-\delta)\Pi^j = (1-\delta)\delta\Pi^{j+1} + (1-\delta)h\left(p^j\right) \geq \delta h\left(p^{j+1}\right) + (1-\delta)h\left(p^j\right) \geq h\left(p^j\right)$ where the first inequality is the induction hypothesis and the second follows from the monotonicity of h. Consequently $(1-\delta)\Pi^k \geq h(p^k)$ for all $k \in \{1, \ldots, n\}$. As $\Pi^{k+1} - \Pi^k = (1-\delta)\Pi^{k+1} - h(p^k) \geq h\left(p^{k+1}\right) - h\left(p^k\right) > 0$, it follows that Π^k is strictly increasing in k.

For existence of a non-babbling Markov equilibrium, a $\beta^k \in [0,1]$ has to exist to make E indifferent between the two recommendations in case option 2 is the bonus option. For $\beta^k = 1$, E strictly prefers to recommend option 2. As the incentives to recommend option 1 are strictly decreasing in β^k, a $\beta^k \in [0,1]$ will exist if and only if E prefers recommending option 1 (in case option 2 is the bonus option) for $\beta^k = 0$. That is, if

$$\begin{aligned}1 + (1-p^k)\delta\Pi^{k+1} &\leq p^k\delta\Pi^{k+1} \\ \Leftrightarrow \Pi^{k+1} &\geq \frac{1}{\delta(2p^k-1)}.\end{aligned}$$

This condition is most demanding for $k = 1$ because p^k and Π^k are both increasing in k. Hence, a non-babbling Markov equilibrium exists if and only if $\Pi^2 \geq 1/(\delta(2p^1-1))$. Plugging in the above derived expression for Π^2, this is condition (7). □

Proof of Lemma 3. The first claim is proven by a simple strategy copying argument. To show the monotonicity of Π_0^k in k let α_t^k be E's best response strategy to C's m-strategy. More precisely, α_t^k is the probability with which E recommends option 1 when it is not the bonus option (after t misses when the signal precision is p^k). To show that $\Pi_0^{k+1} \geq \Pi_0^k$, we will show that E can achieve a value of Π_0^k at precision $k+1$ (after 0 misses). Note that a signal of precision p^{k+1} is sufficient for a signal of precision p^k. That is, E could inject noise into his signal at precision p^{k+1} in order to end up with a signal of precision p^k. Suppose for all $\tilde{k} \geq k+1$ (and all t) E injects noise into his signal such that the new signal has precision $p^{\tilde{k}-1}$ and then plays the strategy $\hat{\alpha}_t^{\tilde{k}} = \alpha_t^{\tilde{k}-1}$.

Equivalently, E can use his improved signal and adjust his behavior to inject some noise in this way.[13] Clearly, this will yield a value of Π_0^k (at precision $k+1$ after 0 misses). Hence, Π_0^{k+1} has to be at least Π_0^k (and is usually higher as the described strategy is not optimal).

Next we show an intermediate result: $\Pi_t^k \leq \Pi_0^{k+1}$ in every m-equilibrium. To see this, note that E's payoffs are bounded from above by $1/(1-\delta)$, i.e. the value of recommending the bonus option each period and C never stopping the game. Put differently, per period payoffs are below 1. This implies $(1-\delta)\Pi_0^{k+1} \leq 1$. Now suppose, by way of contradiction, $\Pi_t^k > \Pi_0^{k+1}$. Then also $\Pi_{t-1}^k > \Pi_0^{k+1}$ because E can after $t-1$ misses simply recommend his bonus option which would then give him a value of $1 + \delta\Pi_t^k/2 + \delta\Pi_0^{k+1}/2 > 1 + \delta\Pi_0^{k+1} \geq \Pi_0^{k+1}$ where the first inequality uses $\Pi_t^k > \Pi_0^{k+1}$ and the second inequality uses $(1-\delta)\Pi_0^{k+1} \leq 1$. Hence, $\Pi_{t-1}^k > \Pi_0^{k+1}$. Iterating this argument yields $\Pi_0^k > \Pi_0^{k+1}$. However, $\Pi_0^k > \Pi_0^{k+1}$ contradicts the first result of Lemma 3 shown above. Hence, $\Pi_t^k \leq \Pi_0^{k+1}$ holds in every m-equilibrium.

$\Pi_t^k \geq \Pi_{t+1}^k$ is shown using the intermediate result of the previous paragraph. Let E recommend his bonus option after t misses (at precision k). This (possibly non-optimal strategy) yields a value of $1 + \delta\Pi_0^{k+1}/2 + \delta\Pi_{t+1}^k/2 \geq 1 + \delta\Pi_{t+1}^k \geq \Pi_{t+1}^k$ where the first inequality uses $\Pi_{t+1}^k \leq \Pi_0^{k+1}$ (see previous paragraph) and the second inequality uses $\Pi_{t+1}^k \leq 1/(1-\delta)$. As recommending E's bonus option after t misses yields a value of at least Π_{t+1}^k, the result $\Pi_t^k \geq \Pi_{t+1}^k$ follows. □

Proof of Lemma 4. l^k is the smallest natural number t such that after t consecutive misses, it is optimal for the expert to recommend option 1. We can thus take condition (9) and replace $t-1$ by t. Then, it can be written as $\Pi_{t+1}^k \leq \Pi_0^{k+1} - \frac{1}{(2p^k-1)\delta}$. Hence l^k is the smallest natural number t for which the latter condition holds. More explicitely,

$$\begin{aligned} l^k &= \min\{t \in \mathbb{N} | \Pi_{t+1}^k \leq \Pi_0^{k+1} - \frac{1}{(2p^k-1)\delta}\} \\ &= \max\left\{ 0, \min\{t \in \mathbb{N} | \underbrace{\Pi_t^k \leq \Pi_0^{k+1} - \frac{1}{(2p^k-1)\delta}}_{(*)}\} - 1 \right\}. \end{aligned}$$

[13] More precisely, let $\gamma^{\tilde{k}} = (p^{\tilde{k}} - p^{\tilde{k}-1})/(p^{\tilde{k}} - 1/2)$. This is chosen such that drawing from the prior with probability $\gamma^{\tilde{k}}$ and with the counter probability from a signal technology with precision $p^{\tilde{k}}$ yields a signal of precision $p^{\tilde{k}-1}$. If option 1 is the bonus option, let E recommend option 1 with probability $1-\gamma^{\tilde{k}}/2$ and option 2 with probability $\gamma^{\tilde{k}}/2$. If option 2 is the bonus option, let $\hat{\alpha}_t^{\tilde{k}} = (1-\gamma^{\tilde{k}}/2)\alpha_t^{\tilde{k}-1} + (\gamma^{\tilde{k}}/2)(1-\alpha_t^{\tilde{k}-1})$. This yields $\hat{\Pi}_t^{\tilde{k}} = \Pi_t^{\tilde{k}-1}$.

Using formula (10), we can reformulate the condition $(*)$ via

$$\begin{aligned}(*) &\Leftrightarrow \sum_{j=0}^{m-1-t} (\delta(1-p^k))^j \left(\frac{1}{2} + p^k \delta \Pi_0^{k+1}\right) \leq \Pi_0^{k+1} - \frac{1}{(2p^k-1)\delta} \\ &\Leftrightarrow \sum_{j=0}^{m-1-t} (\delta(1-p^k))^j \leq \underbrace{\frac{\Pi_0^{k+1} - \frac{1}{(2p^k-1)\delta}}{\frac{1}{2} + p^k\delta\Pi_0^{k+1}}}_{=:P} \\ &\Leftrightarrow \frac{1-(\delta(1-p^k))^{m-t}}{1-\delta(1-p^k)} \leq P \\ &\Leftrightarrow \underbrace{1-(1-\delta(1-p^k))P}_{=:A} \leq (\underbrace{\delta(1-p^k)}_{=:B})^{m-t}\end{aligned}$$

Looking at this last inequality, we see that it is always satisfied if $A \leq 0$ and that it is never satisfied if $P < 0$ (which is equivalent to $\Pi_0^{k+1} < \frac{1}{(2p^k-1)\delta}$). These cases correspond to $l^k = 0$ and $l^k = m$, respectively. In all the other cases, we can apply the natural logarithm on both sides since they will be positive. We continue:

$$\begin{aligned}&\Rightarrow \ln(A) \leq (m-t)\ln(B) \\ &\Leftrightarrow \frac{\ln(A)}{\ln(B)} \geq m-t \qquad \text{(as } \ln(B) < 0) \\ &\Leftrightarrow t \geq m - \frac{\ln(A)}{\ln(B)}\end{aligned}$$

This implies that $l^k = \left\lceil m - 1 - \frac{\ln(A)}{\ln(B)} \right\rceil$ whenever the number inside the ceiling function is larger than -1 and $l^k = 0$ else. □

Proof of Lemma 5. Define $\tilde{\Pi}_t(\Pi_0)$ (for $t \in \{0, 1, \ldots, m\}$) by iterating backwards starting from $\tilde{\Pi}_m = 0$ and using the following formula:

$$\tilde{\Pi}_{t-1}(\Pi_0) = \begin{cases} 1/2 + p^n\delta\Pi_0 + (1-p^n)\delta\tilde{\Pi}_t(\Pi_0) & \text{if } \Pi_0 - \tilde{\Pi}_t(\Pi_0) \geq \frac{1}{\delta(2p^n-1)} \\ 1 + \frac{1}{2}\delta\Pi_0 + \frac{1}{2}\delta\tilde{\Pi}_t(\Pi_0) & \text{else.} \end{cases} \tag{17}$$

Note that the case distinction is done such that $\tilde{\Pi}_t$ is continuous ($\tilde{\Pi}_{t-1}(\Pi_0)$ is simply the maximum of the expression in the first and second case). Clearly, the derivative $\tilde{\Pi}'_{m-1}$ exists for almost all values of Π_0 and is in $\{\delta/2, p^n\delta\}$. Hence, $\tilde{\Pi}'_{m-1} < \delta \leq 1$. Iterating backwards, $\tilde{\Pi}_t$ is continuous and its derivative exists for almost all Π_0. Furthermore, $\tilde{\Pi}'_{t-1}(\Pi_0) \in \{p^n\delta + (1-p^n)\delta\tilde{\Pi}'_t(\Pi_0), \delta/2 + \delta\tilde{\Pi}'_t(\Pi_0)/2\}$ and therefore – given that $\tilde{\Pi}'_t < \delta$ – we have $\tilde{\Pi}'_{t-1} < \delta$. In particular, $\tilde{\Pi}_0(\Pi_0)$ is continuous and has a derivative (which exists almost everywhere) that is strictly positive and strictly smaller than $\delta \leq 1$. The operator $\tilde{\Pi}_0$ is therefore a contraction and the equation $\tilde{\Pi}_0(\Pi_0) = \Pi_0$ has a unique solution Π_0^* by the contraction mapping theorem.

Next we show that $\Pi_0^* \in (0, 1/(1-\delta))$. To this purpose it is sufficient to show $\tilde{\Pi}_0(0) > 0$ and $\tilde{\Pi}_0(1/(1-\delta)) < 1/(1-\delta)$. Clearly, $\tilde{\Pi}_{t-1}(0) > 0$ and in particular $\tilde{\Pi}_0(0) > 0$ holds. Turning to $\tilde{\Pi}_0(1/(1-\delta)) < 1/(1-\delta)$, note that $\tilde{\Pi}_{m-1}(1/(1-\delta)) < 1/(1-\delta)$ as both $1/2 + p^n\delta/(1-\delta) < 1/(1-\delta)$ and $1 + \delta/(2(1-\delta)) < 1/(1-\delta)$ hold. Now proceeding by backward induction $\tilde{\Pi}_{t-1}((1/(1-\delta)) < 1/(1-\delta)$ given that $\tilde{\Pi}_t((1/(1-\delta)) < 1/(1-\delta)$ as both $1/2 + p^n\delta/(1-\delta) + (1-p^n)\delta/(1-\delta) \leq 1/(1-\delta)$ and $1 + \delta/(2(1-\delta)) + \delta/(2(1-\delta)) \leq 1/(1-\delta)$ hold. Given that $\Pi_0^* \in (0, 1/(1-\delta))$, also $\tilde{\Pi}_{t-1}(\Pi_0^*) \in (0, 1/(1-\delta))$ for all $t \in \{1, \ldots, m\}$ by the same steps.

Note that E's value when playing best response against an m-strategy has to satisfy $\Pi_t^n = \tilde{\Pi}_t(\Pi_0^n)$ for all $t \in \{0, \ldots, m-1\}$. As we have just shown, there exists a unique solution to this condition and this solution is feasible, i.e. E's value is in $(0, 1/(1-\delta))$. E's best response strategy is given by the case distinctions in (17): If $\Pi_0^* - \tilde{\Pi}_t(\Pi_0^*) \geq 1/(\delta(2p^n - 1))$, then E recommends option 1 after $t-1$ misses. Otherwise, E recommends his bonus option. Finally, note that E's best response is a cutoff strategy as $\tilde{\Pi}_t(\Pi_0^*)$ is decreasing in t. This can be shown as in the proof of Lemma 3. □

Proof of Proposition 3. By (1), an m-equilibrium cannot exist if E always recommends his bonus option. This strategy yields a payoff after 0 misses of

$$\Pi_0 = \sum_{j=0}^{m-1} \left(\frac{\delta}{2}\right)^j (1 + \delta\Pi_0/2) = \frac{(1 + \delta\Pi_0/2)(1 - (\delta/2)^m)}{1 - \delta/2}$$

which can be solved for Π_0 yielding

$$\Pi_0^* = \frac{1 - (\delta/2)^m}{1 - \delta + (\delta/2)^{m+1}}.$$

Always recommending the bonus option is not E's best response if after $m-1$ misses (9) holds with Π_0^* in place of Π_0^{k+1} and zero in place of Π_{t+1}^k, i.e. if

$$\frac{1 - (\delta/2)^m}{1 - \delta + (\delta/2)^{m+1}} \geq \frac{1}{\delta(2p^n - 1)}.$$

If the opposite of this inequality holds, then always recommending the bonus option is E's best response to C's m-strategy (and this best response is unique by Lemma 5) and therefore no m-equilibrium can exist. This gives the condition in (13). If (13) does not hold, then E's unique best response to C's m-strategy includes recommending option 1 after m-1 misses. This implies that C's value when using his m-strategy is strictly above $(1/2)/(1-\delta)$ (given that V_O satisfies (1)) and therefore there exist values of $V_O > (1/2)/(1-\delta)$ such that C's value is above V_O if C plays an m-strategy and E plays his best response to this strategy. □

Proof of Proposition 4. We compute E's conditional value depending on the event $w \in \mathcal{W}$ that occured and then sum over all possible events (making a distinction between histories in which learning level n is reached and those in which the advice relationship was dissolved before). More concretely, we get

$$\Pi_0 = \sum_{\bar{w} \in \mathcal{W}} \mathbb{P}(w = \bar{w}) \mathbb{E}(\Pi_0 | w = \bar{w}).$$

For $\bar{w} \in \mathcal{W}_n$, we get

$$\mathbb{P}(w = \bar{w}) = \prod_{i=1}^{n-1} \mathbb{P}(w_i = \bar{w}_i) = \prod_{i=1}^{n-1} \left(\mathbb{1}_{\{\bar{w}_i \leq l^i\}} (\frac{1}{2})^{\bar{w}_i} + \mathbb{1}_{\{\bar{w}_i > l^i\}} (\frac{1}{2})^{l^i} (1 - p^i)^{\bar{w}_i - l^i - 1} p^i \right),$$

$$\mathbb{E}(\Pi_0 | w = \bar{w}) = \sum_{k=1}^{n-1} H_k \delta^{\sum_{j=1}^{k-1} \bar{w}_j} + \delta^{\sum_{j=1}^{n-1} \bar{w}_j} C_{n-1}.$$

In the above equation, H_k denotes the expert's expected value in learning level p^k (in which he will spend $\bar{w}_k$ periods). Moreover, C_{n-1} is the expert's continuation value after the $(n-1)-$th hit at the first period with learning level p^n. Both H_k and C_{n-1} are computed below.

$$H_k = \mathbb{1}_{\{\bar{w}_k \leq l^k\}} \sum_{h=0}^{\bar{w}_k - 1} \delta^h + \mathbb{1}_{\{\bar{w}_k > l^k\}} \left(\sum_{h=0}^{l^k - 1} \delta^h + \frac{1}{2} \sum_{h=l^k}^{\bar{w}_k - 1} \delta^h \right) \tag{18}$$

$$C_{n-1} = \sum_{g=0}^{l^n - 1} (\frac{\delta}{2})^g + (\frac{\delta}{2})^{g+1} C_{n-1} + \sum_{g=l^n}^{m-1} (\frac{1}{2})^{l^n} \delta^g (1 - p^n)^{g - l^n} p^n C_{n-1} + \delta^g (\frac{1}{2})^{g - l^n} \tag{19}$$

$$\Leftrightarrow C_{n-1} = \frac{\sum_{g=0}^{l^n - 1} (\frac{\delta}{2})^g + \sum_{g=l^n}^{m-1} \delta^g (\frac{1}{2})^{g - l^n}}{1 - \sum_{g=0}^{l^n - 1} (\frac{\delta}{2})^{g+1} - \sum_{g=l^n}^{m-1} (\frac{1}{2})^{l^n} \delta^g (1 - p^n)^{g - l^n} p^n}$$

In the above computations, (18) follows since the expert will recommend his bonus option l^k times after reaching a new learning level (assuming that all these recommendations produce misses). Only after l^k misses, he will recommend option 1, which yields him $\frac{1}{2}$ per period in expectation, since bonus option and option 1 are drawn independently. Equation (19) reflects the fact that the experts continuation value after $n-1$ hits and after n hits (or more) is the same, since no further learning happens after precision p^n is reached.

For $\bar{w} \in \mathcal{W}_f$, we get

$$\mathbb{P}(w = \bar{w}) = \prod_{i=1}^{len(\bar{w})} \mathbb{P}(w_i = \bar{w}_i)$$

$$= (\frac{1}{2})^{l^{len(\bar{w})}} (1 - p^{len(\bar{w})})^{m - l^{len(\bar{w})}} \prod_{i=1}^{len(\bar{w})-1} \left(\mathbb{1}_{\{\bar{w}_i \leq l^i\}} (\frac{1}{2})^{\bar{w}_i} + \mathbb{1}_{\{\bar{w}_i > l^i\}} (\frac{1}{2})^{l^i} (1 - p^i)^{\bar{w}_i - l^i - 1} p^i \right), \tag{20}$$

$$\mathbb{E}(\Pi_0 | w = \bar{w}) = \sum_{k=1}^{len(\bar{w})-1} H_k \delta^{\sum_{j=1}^{k-1} \bar{w}_j} + \delta^{\sum_{j=1}^{len(\bar{w})-1} \bar{w}_j} \left(\sum_{h=0}^{l^{len(\bar{w})}-1} \delta^h + \frac{1}{2} \sum_{h=l^{len(\bar{w})}}^{m-1} \delta^h \right). \tag{21}$$

Equation (20) follows since m consecutive misses in learning level $len(\bar{w})$ only occur if the bonus option was different from option 1 for $l^{len(\bar{w})}$ periods in a row and the expert failed to generate good advice in the remaining $m - l^{len(\bar{w})}$ periods. Likewise, the term in the brackets in (21) describes the expected payoff of the expert in the learning level in which m consecutive misses are produced. Putting all the above formulas together yields the desired result. □

Proof of Proposition 5. The probability of having m consecutive misses conditional on reaching precision level k does of course not only depend on the precision level, but also on the strategy l^k of the expert. More concretely,

$$HR(k) = \left(\frac{1}{2}\right)^{l^k} (1 - p^k)^{m - l^k}.$$

As $1 - p^k < \frac{1}{2}$ for all k by assumption, for (weakly) decreasing values of l (i.e. $l^{k+1} \leq l^k$) the hazard rate $HR(k)$ is (strictly) decreasing in k. When l^k is strictly increasing, then it cannot increase to $l^{k+1} = m$ by Remark 1. Hence, we always have $m - l^{k+1} > 0$ in this case. We now derive the sufficient condition for the hazard rate to be (weakly) decreasing:

$$\begin{aligned} HR(k+1) \leq HR(k) &\Leftrightarrow \left(\frac{1}{2}\right)^{l^{k+1}} (1 - p^{k+1})^{m - l^{k+1}} \leq \left(\frac{1}{2}\right)^{l^k} (1 - p^k)^{m - l^k} \\ &\Leftrightarrow \left(\frac{1}{2(1 - p^k)}\right)^{l^{k+1} - l^k} \leq \left(\frac{1 - p^k}{1 - p^{k+1}}\right)^{m - l^{k+1}} \end{aligned}$$

Since the LHS of the last inequality is increasing in l^{k+1} and decreasing in l^k and the RHS is decreasing in l^{k+1}, it is sufficient to consider $l^{k+1} = m - 1$ and $l^k = 0$ (the extreme cases), which yields

$$\left(\frac{1}{2(1-p^k)}\right)^{m-1} \leq \frac{1-p^k}{1-p^{k+1}} \Leftrightarrow 1-p^{k+1} \leq (1-p^k)^m 2^{m-1}$$
$$\Leftrightarrow p^{k+1} \geq 1-(1-p^k)^m 2^{m-1}$$

This concludes the proof. □

Proof of Proposition 6. We will show by induction that $l^k = 1$ implies $l^{k-1} = 1$ in an $m = 1$ equilibrium. However, note first that $l^n = 0$ in an $m = 1$ equilibrium as the consumer would otherwise be better off by ending the advice relationship immediately when reaching precision level p^n.

Now assume that $l^k = 1$ for some $k \in \{2, \ldots, n-1\}$. This implies that the expected payoff of the expert when choosing $l^k = 1$, namely $1 + \delta\Pi_0^{k+1}/2$, is greater or equal than his expected payoff when choosing $l^k = 0$, namely $1/2 + p^k\delta\Pi_0^{k+1}$. Put differently, $1+\delta\Pi_0^{k+1}/2 \geq 1/2+p^k\delta\Pi_0^{k+1}$ or equivalently $0 \geq -1/2+\delta\Pi_0^{k+1}(p^k-1/2)$. As $0 < \Pi_0^k \leq \Pi_0^{k+1}$ and $1/2 \leq p^{k-1} < p^k$ this inequality implies $0 \geq -1/2 + \delta\Pi_0^k(p^{k-1} - 1/2)$ which is equivalent to saying that the expected payoff of the expert is higher when choosing $l^{k-1} = 1$ than when choosing $l^{k-1} = 0$. Consequently, $l^k = 1$ implies $l^{k-1} = 1$.

Hence, in an $m = 1$ equilibrium $l^k = 1$ for $k \leq \bar{k}$ and $l^k = 0$ for $k > \bar{k}$ for some $\bar{k} \in \{0, \ldots, n\}$. The result on V_0^k now readily follows as an increase in k improves the quality of advice in two ways: (i) l^k may decrease and, (ii) p^k increases.

More formally, $V_0^n = p^n(1 + \delta V_0^n) \Leftrightarrow V_0^n = p^n/(1 - p^n\delta)$ and $V_0^k = p^k(1 + \delta V_0^{k+1})$ for $k \in \{\bar{k}+1, \ldots, n-1\}$. For now let $\bar{k} \leq n-2$, then $V_0^n > V_0^{n-1}$ holds as $p^n/(1-\delta p^n) > p^{n-1}\left(1+\delta p^n/(1-\delta p^n)\right) \Leftrightarrow p^{n-1}/p^n < (1-\delta p^n)/(1-\delta p^n) = 1$ which is true by $p^{n-1} < p^n$.[14] Using this as the starting point for backward induction $V_0^k = p^k(1+\delta V_0^{k+1}) > p^{k-1}(1+\delta V_0^k) = V_0^{k-1}$ by the induction hypothesis $V_0^{k+1} > V_0^k$ and $p^k > p^{k-1}$ for all $k-1 > \bar{k}$. The backward induction logic extends to $\bar{k}$ where $V_0^{\bar{k}} = (1+\delta V_0^{\bar{k}+1})/2 < p^{\bar{k}+1}\left(1+\delta V_0^{\bar{k}+2}\right) = V_0^{\bar{k}+1}$ by $1/2 < p^{\bar{k}+1}$ and $V_0^{\bar{k}+1} < V_0^{\bar{k}+2}$. The backward induction argument continues further for $k < \bar{k}$ as there $V_0^k = (1+\delta V_0^{k+1})/2 < (1+\delta V_0^{k+2})/2 = V_0^{k+1}$ where the inequality follows from the induction hypothesis $V_0^{k+1} < V_0^{k+2}$. □

Proof of Lemma 6. We are using formulas (10) and (11) to compute the expert's value Π_0^2 for different values of l^2.

[14]Clearly, the argument below still holds true for the case $\bar{k} = n-1$.

i) $l^2 = 0$

$$\Rightarrow \Pi_0^2 = \frac{1}{2} + p^2 \delta \Pi_0^2$$
$$\Leftrightarrow \Pi_0^2 = \frac{1}{2(1 - p^2\delta)}$$

ii) $l^2 = 1$

$$\Rightarrow \Pi_0^2 = 1 + \frac{1}{2}\delta \Pi_0^2$$
$$\Leftrightarrow \Pi_0^2 = \frac{1}{1 - \frac{1}{2}\delta}$$

This implies that $l^2 = 1$ yields a higher expected value than $l^2 = 0$ for the expert if

$$1 - \frac{1}{2}\delta < 2(1 - p^2\delta)$$
$$\Leftrightarrow \delta < \frac{1}{2p^2 - \frac{1}{2}}$$

holds. The only thing that is left to check now is what the optimal choice for l^1 is in each of the two cases above.

i)

$$l^1 = 0 \Rightarrow \Pi_0^1 = \frac{1}{2} + p^1 \delta \Pi_0^2 = \frac{1}{2} + \frac{p^1\delta}{2(1 - p^2\delta)}$$
$$l^1 = 1 \Rightarrow \Pi_0^1 = 1 + \frac{\delta}{2}\Pi_0^2 = 1 + \frac{\delta}{4(1 - p^2\delta)}$$
$$\frac{1}{2} + \frac{p^1\delta}{2(1 - p^2\delta)} > 1 + \frac{\delta}{4(1 - p^2\delta)} \Leftrightarrow \frac{(p^1 - \frac{1}{2})\delta}{2(1 - p^2\delta)} > \frac{1}{2} \Leftrightarrow \delta > \frac{1}{p^1 + p^2 - \frac{1}{2}}$$

Hence, in the case $l^2 = 0$, the expert will choose $l^1 = 0$ if $\delta > \frac{1}{p^1 + p^2 - \frac{1}{2}}$ and he will choose $l^1 = 1$ if $\delta < \frac{1}{p^1 + p^2 - \frac{1}{2}}$ holds.

ii)

$$l^1 = 0 \Rightarrow \Pi_0^1 = \frac{1}{2} + p^1 \delta \Pi_0^2 = \frac{1}{2} + \frac{p^1\delta}{1 - \frac{1}{2}\delta}$$
$$l^1 = 1 \Rightarrow \Pi_0^1 = 1 + \frac{\delta}{2}\Pi_0^2 = 1 + \frac{\delta}{2 - \delta}$$
$$\frac{1}{2} + \frac{p^1\delta}{1 - \frac{1}{2}\delta} > 1 + \frac{\delta}{2 - \delta} \Leftrightarrow \delta > \frac{1}{2p^1 - \frac{1}{2}}$$

Since the latter equation is never satisfied for $l^2 = 1$ due to $\delta < \frac{1}{2p^2 - \frac{1}{2}} < \frac{1}{2p^1 - \frac{1}{2}}$, $l^2 =$

1 always implies $l^1 = 1$.

This completes the proof. $\square$

Proof of Lemma 7. Solving via backward induction, we start determining l^2 by going through three different cases.

1. $l^2 = 0$: Then $\Pi_1^2 = 1/2 + p^2\delta\Pi_0^2$ and $\Pi_0^2 = 1/2 + p^2\delta\Pi_0^2 + (1-p^2)\delta\Pi_1^2$. Plugging the first expression into the second one and solving for Π_0^2 yields

$$\Pi_0^2 = \frac{(1+\delta)/2 - \delta p^2/2}{1 - p^2\delta - (1-p^2)\delta^2 p^2}. \tag{22}$$

2. $l^2 = 1$: Then $\Pi_1^2 = 1/2 + p^2\delta\Pi_0^2$ and $\Pi_0^2 = 1 + \delta\Pi_0^2/2 + \delta\Pi_1^2/2$ which can be solved for

$$\Pi_0^2 = \frac{1+\delta/4}{1-\delta/2-\delta^2p^2/2}. \tag{23}$$

3. $l^2 = 2$: Then $\Pi_1^2 = 1 + \delta\Pi_0^2/2$ and $\Pi_0^2 = 1 + \delta\Pi_0^2/2 + \delta\Pi_1^2/2$ which can be solved for

$$\Pi_0^2 = \frac{1+\delta/2}{1-\delta/2-\delta^2/4}. \tag{24}$$

Therefore, $l^2 = 1$ is the expert's best response if and only if

$$\frac{1+\delta/4}{1-\delta/2-\delta^2p^2/2} > \max\left\{\frac{1+\delta/2}{1-\delta/2-\delta^2/4}, \frac{(1+\delta)/2-\delta^2p^2/2}{1-p^2\delta-(1-p^2)\delta^2p^2}\right\}.$$

Conditional on $l^2 = 1$ being the expert's best response in learning level 2, we will now check under which conditions $l^1 = 0$ is the expert's best response in learning level 1. Again, we have to go through three cases.

1. $l^1 = 0$: Then $\Pi_1^1 = 1/2 + p^1\delta\Pi_0^2$ and $\Pi_0^1 = 1/2 + p^1\delta\Pi_0^2 + (1-p^1)\delta\Pi_1^1$. Plugging in yields

$$\Pi_0^1 = \frac{1}{2} + \frac{(1-p^1)\delta}{2} + \frac{p^1\delta(1+(1-p^1)\delta)(1+\delta/4)}{1-\delta/2-\delta^2p^2/2}.$$

2. $l^1 = 1$: Then $\Pi_1^1 = 1/2 + p^1\delta\Pi_0^2$ and $\Pi_0^1 = 1 + \delta\Pi_0^2/2 + \delta\Pi_1^1/2$. Plugging in yields

$$\Pi_0^1 = 1 + \delta/4 + \frac{(\delta/2+\delta^2p^1/2)(1+\delta/4)}{1-\delta/2-\delta^2p^2/2}.$$

3. $l^1 = 2$: Then $\Pi_1^1 = 1 + \delta\Pi_0^2/2$ and $\Pi_0^1 = 1 + \delta\Pi_0^2/2 + \delta\Pi_1^1/2$. Plugging in yields

$$\Pi_0^1 = 1 + \delta/2 + \frac{(\delta/2 + \delta^2/4)(1 + \delta/4)}{1 - \delta/2 - \delta^2 p^2/2}.$$

Therefore, $l^1 = 0$ will be the expert's best response if and only if

$$\frac{1}{2} + \frac{(1-p^1)\delta}{2} + \frac{p^1\delta(1 + (1-p^1)\delta)(1+\delta/4)}{1 - \delta/2 - \delta^2 p^2/2} \\ \geq \max\left\{1 + \delta/4 + \frac{(\delta/2 + \delta^2 p^1/2)(1+\delta/4)}{1 - \delta/2 - \delta^2 p^2/2}, 1 + \delta/2 + \frac{(\delta/2 + \delta^2/4)(1+\delta/4)}{1 - \delta/2 - \delta^2 p^2/2}\right\}.$$

□

B. Experiment Instructions

B.1. Control Group

Freiwilligkeit des Experimentes

Die Teilnahme an diesem Experiment ist freiwillig. Sie können die Teilnahme jederzeit ohne Angabe von Gründen abbrechen.

Instruktionen

Bitte lesen Sie die folgenden Instruktionen sorgfältig. Vor dem Experiment bekommen Sie einige Kontrollfragen gestellt und Sie können bei korrekter Beantwortung Geld gewinnen. Konkret werden Ihnen **vier Kontrollfragen** gestellt. Hiervon wird nach Ihren Antworten eine zufällig ausgewählt und wenn Ihre Antwort auf diese Frage beim ersten Versuch richtig war, bekommen Sie eine zusätzliche Auszahlung von **1,00€**.

Im Folgenden werden Sie zufällig in Zweiergruppen eingeteilt und werden mit Ihrem zugeteilten Spielpartner ein Spiel spielen. In diesem Spiel können Sie Spielpunkte erspielen. Auf Basis dieser Spielpunkte wird am Ende Ihre Auszahlung ermittelt, was weiter unten erläutert wird. Zusätzlich erhalten Sie eine hiervon unabhängige Auszahlung von **4,00€** für das Erscheinen und Ihre Teilnahme am Experiment. In dem Spiel werden Sie zufällig entweder die Rolle von Spieler A oder von Spieler B übernehmen. Das Spiel wird nun beschrieben und danach anhand eines Beispiels für zwei Spielrunden veranschaulicht. Dort sehen Sie auch beispielhaft die Bildschirmanzeigen, die beiden Spielern jeweils angezeigt werden.

Entscheidungen der Spieler

Das Spiel wird über mehrere Runden gespielt und in jeder Runde hat Spieler A die Wahl zwischen Option 1 und Option 2 und Spieler B entscheidet in der Folge, ob eine weitere Runde des Spiels gespielt wird. Es gibt in jeder Runde vier mögliche Fälle, die alle gleich wahrscheinlich sind und vor jeder neuen Runde zufällig bestimmt werden.

	Auszahlungen der Spieler bei Wahl von	
Fälle	Option 1	Option 2
1. Fall	A: 1 Punkt, B: 1 Punkt	A: 0 Punkte, B: 0 Punkte
2. Fall	A: 0 Punkte, B: 0 Punkte	A: 1 Punkt, B: 1 Punkt
3. Fall	A: 1 Punkt, B: 0 Punkte	A: 0 Punkte, B: 1 Punkt
4. Fall	A: 0 Punkte, B: 1 Punkt	A: 1 Punkt, B: 0 Punkte

Abbildung 1: Übersicht über die möglichen Auszahlungen für Spieler A und B

In jedem möglichen Fall erhält also jeder Spieler eine Auszahlung von 1 von **genau einer** der beiden Optionen, während die andere Option ihm eine Auszahlung von 0 gibt. Die Option mit der höheren Auszahlung kann entweder für beide Spieler die gleiche oder aber eine unterschiedliche sein.

In jeder Spielrunde tritt genau einer der obigen Fälle ein, aber keiner der Spieler weiß mit Sicherheit, welcher das ist. Spieler A bekommt jedoch immer angezeigt für welche der Optionen er einen Punkt erhält. Darüber hinaus erhält er einen automatisch erzeugten *Hinweis* darüber, welche Option Spieler B einen Punkt einbringen *könnte*. Dieser Hinweis ist immer mit einer **Wahrscheinlichkeit von 82% korrekt** und mit einer **Wahrscheinlichkeit von 18% inkorrekt**.

Nach der Entscheidung von Spieler A werden beide Spieler über ihre daraus resultierenden Auszahlungen informiert. Spieler B erfährt hierbei nur, ob er eine Auszahlung von 1 oder 0 (Spielpunkten) erhält und nicht, was der Hinweis von Spieler A war oder welche Auszahlung Spieler A erhalten hat. Spieler A wird hingegen auch über die Auszahlung von Spieler B informiert. Spieler B kann also keine der Optionen selbst wählen, sondern erhält seine Auszahlung abhängig von der Wahl von Spieler A. Im Anschluss daran kann Spieler B entscheiden, ob er das Spiel *beenden* oder für eine weitere Runde *fortführen* möchte.

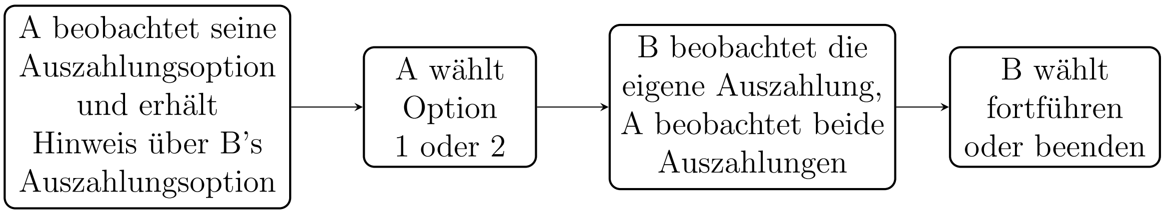
- Spieler B wählt *fortführen*:
  In diesem Fall wird mit einer Wahrscheinlichkeit von 90% eine weitere Runde des Spiels gespielt.

  Mit einer Wahrscheinlichkeit von 10% **endet das Spiel** trotz der Entscheidung von Spieler B das Spiel fortzuführen (sonst könnte das Spiel theoretisch unendlich lange dauern). Beide Spieler erhalten ihre bis dahin erspielten Spielpunkte. Beide Spieler erhalten die Nachricht, dass das Spiel *exogen* beendet wurde.

- Spieler B wählt *beenden*:
  In diesem Fall bekommt Spieler B zusätzlich **5 Spielpunkte** gutgeschrieben, Spieler

A erhält keine weiteren Punkte. Das Spiel ist zu Ende und beide Spieler werden darüber informiert, dass Spieler B das Spiel beendet hat.

Jede Runde des Spiels kann wie folgt in einem Schaubild veranschaulicht werden:



Neues Spiel mit neuem Spielpartner

Sobald ein Spiel für alle Spieler beendet ist (entweder exogen oder weil alle Spieler B ihr Spiel beendet haben), werden die Spielpartner neu zugelost. Jeder behält hierbei jedoch seine Rolle als Spieler A oder Spieler B und bekommt zufällig einen Spieler des anderen Typs zugelost. Das Spiel wird erneut gestartet. Insgesamt werden **10 Spiele** mit wechselnden Spielpartnern durchgeführt. Am Ende wird zufällig **eines** der 10 Spiele ausgewählt und die dort erspielte Punktzahl wird nach Beendigung des Experimentes (zusammen mit der festen Auszahlung) ausgezahlt. Ein Spielpunkt entspricht hierbei **1,00€**.

Beispiel

In dem folgenden Beispiel (siehe Abbildung 1) wählt Spieler A in der ersten Runde Option 1 (oben links im Bild). Im Folgenden werden beide Spieler darüber informiert, dass diese Wahl Spieler B eine Auszahlung von 0 einbringt (zu sehen ist nur der Bildschirm von Spieler B, oben rechts). Spieler A kann so feststellen, dass sein Hinweis über Spieler B in Runde 1 korrekt war, da der Hinweis Option 2 lautete und Option 1 Spieler B eine Auszahlung von 0 einbrachte. Somit hätte Option 2 tatsächlich in einer Auszahlung von 1 für Spieler B resultiert. Spieler B weiß allerdings weder welchen Hinweis Spieler A erhalten hat, noch ob Spieler A diesem Hinweis gefolgt ist.

Im Beispiel entscheidet sich Spieler B für "Spiel fortführen" und es wird eine zweite Runde gespielt. Nun entscheidet sich Spieler A für Option 2 (Bild unten links). Diese Wahl führt zu einer Auszahlung von 1 für Spieler B (siehe Bild unten rechts). Spieler B kann nun wieder entscheiden, ob er das Spiel fortführen oder beenden möchte.

Ende des Experimentes

Zum Ende des Experimentes bekommen Sie noch ein paar Fragen gestellt, bei denen Sie teilweise Geld gewinnen können (dies ist dann jeweils vor Beantwortung der Fragen erklärt). Zuletzt geben Sie über ein Formular Ihre Auszahlungsdaten ein, die von der Universität zur Tätigung der Zahlung benötigt werden.

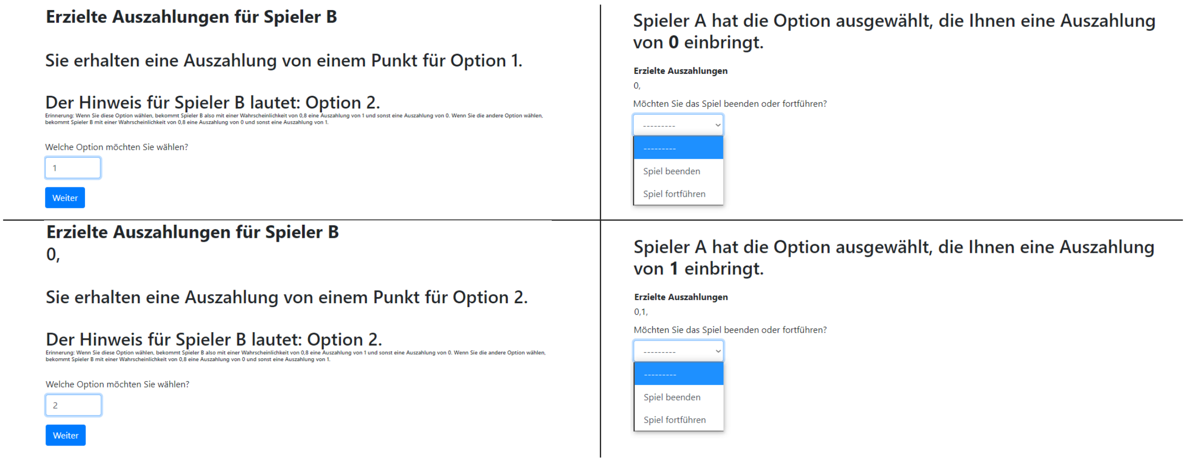


Abbildung 2: Ein Beispiel für die ersten zwei Spielrunden

B.2. Treatment Group

Freiwilligkeit des Experimentes

Die Teilnahme an diesem Experiment ist freiwillig. Sie können die Teilnahme jederzeit ohne Angabe von Gründen abbrechen.

Instruktionen

Bitte lesen Sie die folgenden Instruktionen sorgfältig. Vor dem Experiment bekommen Sie einige Kontrollfragen gestellt und Sie können bei korrekter Beantwortung Geld gewinnen. Konkret werden Ihnen **fünf Kontrollfragen** gestellt. Hiervon wird nach Ihren Antworten eine zufällig ausgewählt und wenn Ihre Antwort auf diese Frage beim ersten Versuch richtig war, bekommen Sie eine zusätzliche Auszahlung von **1,00€**.

Im Folgenden werden Sie zufällig in Zweiergruppen eingeteilt und werden mit Ihrem zugeteilten Spielpartner ein Spiel spielen. In diesem Spiel können Sie Spielpunkte erspielen. Auf Basis dieser Spielpunkte wird am Ende Ihre Auszahlung ermittelt, was weiter unten erläutert wird. Zusätzlich erhalten Sie eine hiervon unabhängige Auszahlung von **4,00€** für das Erscheinen und Ihre Teilnahme am Experiment. In dem Spiel werden Sie zufällig entweder die Rolle von Spieler A oder von Spieler B übernehmen. Das Spiel wird nun beschrieben und danach anhand eines Beispiels für zwei Spielrunden veranschaulicht. Dort sehen Sie auch beispielhaft die Bildschirmanzeigen, die beiden Spielern jeweils angezeigt werden.

Entscheidungen der Spieler

Das Spiel wird über mehrere Runden gespielt und in jeder Runde hat Spieler A die Wahl zwischen Option 1 und Option 2 und Spieler B entscheidet in der Folge, ob eine weitere Runde des Spiels gespielt wird. Es gibt in jeder Runde vier mögliche Fälle, die alle gleich wahrscheinlich sind und vor jeder neuen Runde zufällig bestimmt werden.

	Auszahlungen der Spieler bei Wahl von	
Fälle	Option 1	Option 2
1. Fall	A: 1 Punkt, B: 1 Punkt	A: 0 Punkte, B: 0 Punkte
2. Fall	A: 0 Punkte, B: 0 Punkte	A: 1 Punkt, B: 1 Punkt
3. Fall	A: 1 Punkt, B: 0 Punkte	A: 0 Punkte, B: 1 Punkt
4. Fall	A: 0 Punkte, B: 1 Punkt	A: 1 Punkt, B: 0 Punkte

Abbildung 1: Übersicht über die möglichen Auszahlungen für Spieler A und B

In jedem möglichen Fall erhält also jeder Spieler eine Auszahlung von 1 von **genau einer** der beiden Optionen, während die andere Option ihm eine Auszahlung von 0 gibt. Die Option mit der höheren Auszahlung kann entweder für beide Spieler die gleiche oder aber eine unterschiedliche sein.

In jeder Spielrunde tritt genau einer der obigen Fälle ein, aber keiner der Spieler weiß mit Sicherheit, welcher das ist. Spieler A bekommt jedoch immer angezeigt für welche der Optionen er einen Punkt erhält. Darüber hinaus erhält er einen automatisch erzeugten *Hinweis* darüber, welche Option Spieler B einen Punkt einbringen *könnte*. Dieser Hinweis ist in der ersten Runde mit einer **Wahrscheinlichkeit von 82% korrekt** und mit einer **Wahrscheinlichkeit von 18% inkorrekt**. Die Wahrscheinlichkeit, mit der der Hinweis korrekt ist, nennen wir in dem Experiment die **Hinweisstärke**. Sie wird immer als Dezimalzahl angegeben. Eine Hinweisstärke von 0,82 entspricht zum Beispiel einer Wahrscheinlichkeit von 82%, eine Hinweisstärke von 0,84 entspricht 84%, usw.

Nach der Entscheidung von Spieler A werden beide Spieler über ihre daraus resultierenden Auszahlungen informiert. Spieler B erfährt hierbei nur, ob er eine Auszahlung von 1 oder 0 (Spielpunkten) erhält und nicht, was der Hinweis von Spieler A war oder welche Auszahlung Spieler A erhalten hat. Spieler A wird hingegen auch über die Auszahlung von Spieler B informiert. Spieler B kann also keine der Optionen selbst wählen, sondern erhält seine Auszahlung abhängig von der Wahl von Spieler A. Im Anschluss daran kann Spieler B entscheiden, ob er das Spiel *beenden* oder für eine weitere Runde *fortführen* möchte.

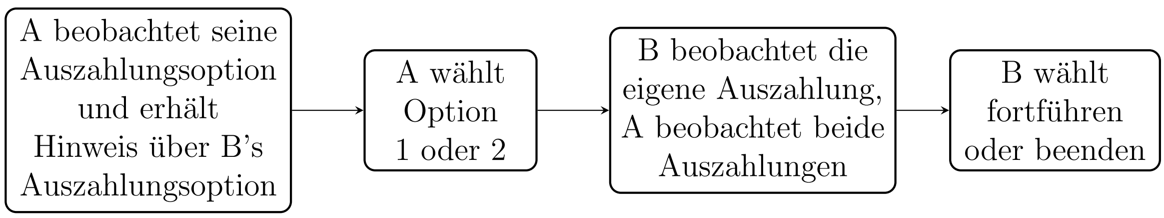
- Spieler B wählt *fortführen*:
  In diesem Fall wird mit einer Wahrscheinlichkeit von 90% eine weitere Runde des Spiels gespielt. Falls Spieler B in der aktuellen Runde eine Auszahlung von einem Spielpunkt erhalten hat, wird in den folgenden Runden der Hinweis, den Spieler A erhält, *verbessert*: Die Wahrscheinlichkeit, mit der der Hinweis korrekt ist, **erhöht sich um 2%** (die Hinweisstärke erhöht sich also um 0,02). Falls Spieler B in der aktuellen Runde eine Auszahlung von null Spielpunkten erhalten hat, bleibt die Hinweisstärke genau wie in der vorherigen Runde.

  Mit einer Wahrscheinlichkeit von 10% **endet das Spiel** trotz der Entscheidung von Spieler B das Spiel fortzuführen (sonst könnte das Spiel theoretisch unendlich lange dauern). Beide Spieler erhalten ihre bis dahin erspielten Spielpunkte. Beide Spieler erhalten die Nachricht, dass das Spiel *exogen* beendet wurde.

- Spieler B wählt *beenden*:
  In diesem Fall bekommt Spieler B zusätzlich **5 Spielpunkte** gutgeschrieben, Spieler A erhält keine weiteren Punkte. Das Spiel ist zu Ende und beide Spieler werden darüber informiert, dass Spieler B das Spiel beendet hat.

Sofern das Spiel über mehrere Runden fortgeführt wird, verbessert sich der Hinweis auch in folgenden Runden (sofern Spieler B eine Auszahlung von einem Punkt erhält). Hierbei erhöht sich die Wahrscheinlichkeit, dass der Hinweis definitiv korrekt ist jeweils um **2%**. Die maximale Wahrscheinlichkeit ist jedoch **90%**. Sollte in einer Runde also diese Wahrscheinlichkeit erreicht sein und Spieler B erhält in dieser Runde nochmals eine Auszahlung von 1, so bleibt die Wahrscheinlichkeit auch in allen folgenden Runden bei 90%.

Jede Runde des Spiels kann wie folgt in einem Schaubild veranschaulicht werden:



Neues Spiel mit neuem Spielpartner

Sobald ein Spiel für alle Spieler beendet ist (entweder exogen oder weil alle Spieler B ihr Spiel beendet haben), werden die Spielpartner neu zugelost. Jeder behält hierbei jedoch seine Rolle als Spieler A oder Spieler B und bekommt zufällig einen Spieler des anderen Typs zugelost. Das Spiel wird erneut gestartet. Insgesamt werden **10 Spiele** mit wechselnden Spielpartnern durchgeführt. Am Ende wird zufällig **eines** der 10 Spiele ausgewählt und die dort erspielte Punktzahl wird nach Beendigung des Experimentes (zusammen mit der festen Auszahlung) ausgezahlt. Ein Spielpunkt entspricht hierbei **1,00€**.

Beispiel

In dem folgenden Beispiel (siehe Abbildung 1) wählt Spieler A in der ersten Runde Option 1 (oben links im Bild). Im Folgenden werden beide Spieler darüber informiert, dass diese Wahl Spieler B eine Auszahlung von 0 einbringt (zu sehen ist nur der Bildschirm von Spieler B, oben rechts). Spieler A kann so feststellen, dass sein Hinweis über Spieler B in Runde 1 korrekt war, da der Hinweis Option 2 lautete und Option 1 Spieler B eine

Auszahlung von 0 einbrachte. Somit hätte Option 2 tatsächlich in einer Auszahlung von 1 für Spieler B resultiert. Spieler B weiß allerdings weder welchen Hinweis Spieler A erhalten hat, noch ob Spieler A diesem Hinweis gefolgt ist.

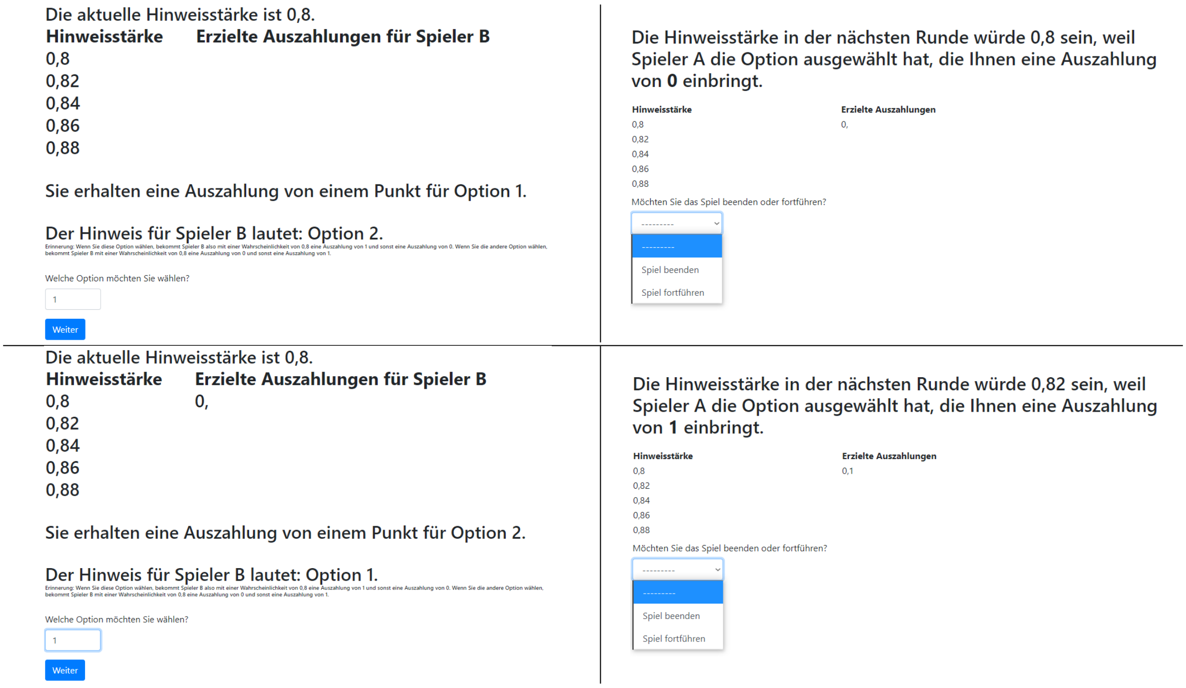


Abbildung 2: Ein Beispiel für die ersten zwei Spielrunden

Im Beispiel entscheidet sich Spieler B für "Spiel fortführen" und es wird eine zweite Runde gespielt. In der zweiten Runde ist die Hinweisstärke dann wiederum 0,82, da Spieler B in der ersten Runde eine Auszahlung von 0 erreicht hat. Nun entscheidet sich Spieler A für Option 2 (Bild unten links). Diese Wahl führt zu einer Auszahlung von 1 für Spieler B (siehe Bild unten rechts). Spieler B kann nun wieder entscheiden, ob er das Spiel fortführen oder beenden möchte und wird darüber informiert, dass die Hinweisstärke in der nächsten Runde 0,84 wäre. Hätte Spieler B in der zweiten Runde eine Auszahlung von 0 erhalten, so wäre die Hinweisstärke in der nächsten Runde weiterhin bei 0,82 geblieben. Die Hinweisstärke erhöht sich immer nur dann, wenn Spieler B in einer Runde eine Auszahlung von 1 erhält.

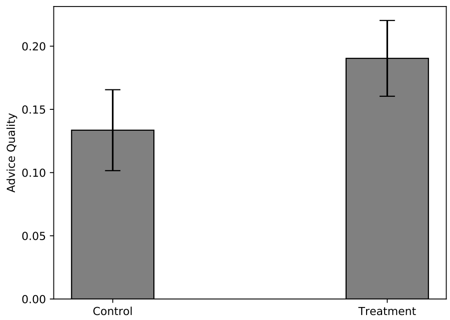
Ende des Experimentes

Zum Ende des Experimentes bekommen Sie noch ein paar Fragen gestellt, bei denen Sie teilweise Geld gewinnen können (dies ist dann jeweils vor Beantwortung der Fragen erklärt). Zuletzt geben Sie über ein Formular Ihre Auszahlungsdaten ein, die von der Universität zur Tätigung der Zahlung benötigt werden.

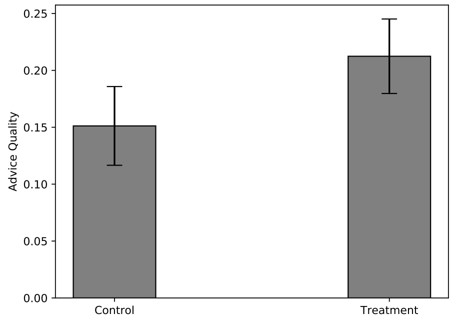
C. Additional Results and Robustness Checks

In this section, we will provide additional results and robustness checks related to the experimental results given in Section 6. We will have a further look at advice quality, welfare distribution and hazard rates in turn.

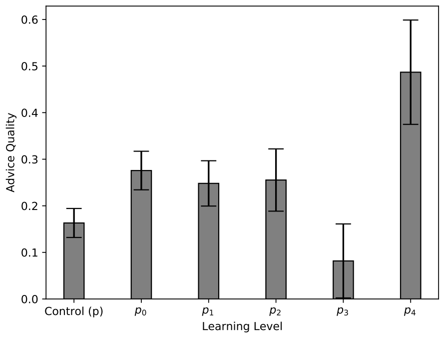
C.1. Advice Quality



(a) Consumers with max. 1 incorrect answer



(b) First two supergames taken out



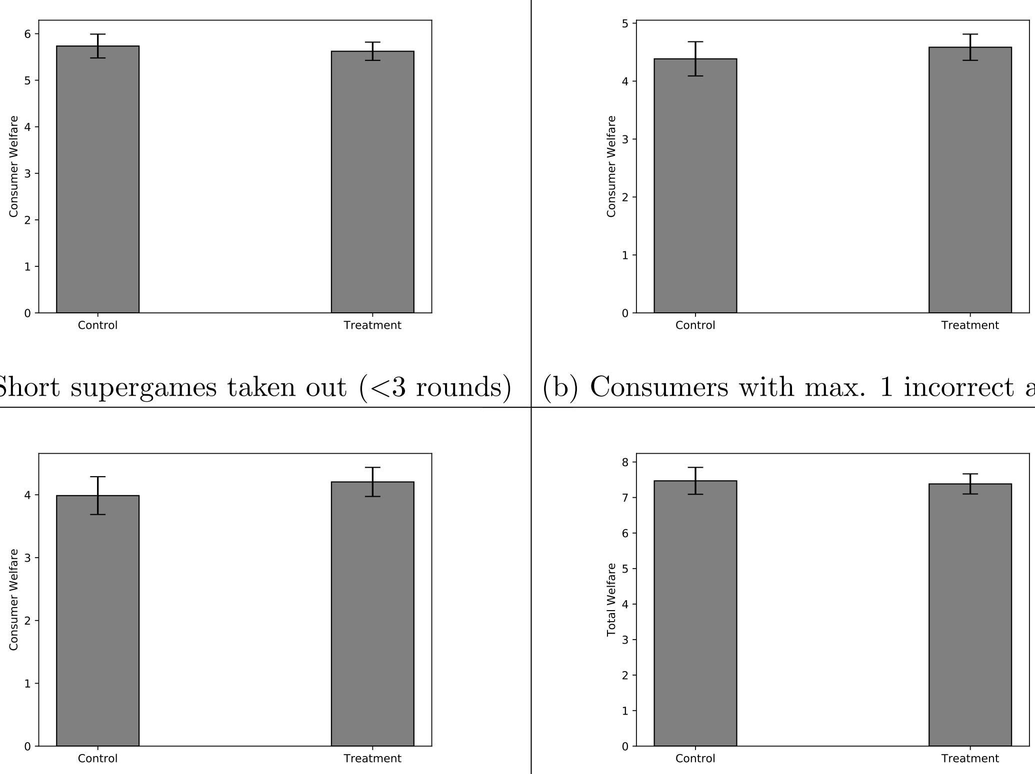
(c) Advice Quality per Learning Level

Figure 7: Robustness Checks for Advice Quality

Figure 7 shows the robustness checks for advice quality. In Figure 7a, we only considered those experts who gave at most one incorrect answer to the check questions. We can see that the advice quality in the treatment group is still higher, but the difference becomes a bit less significant. The same happens when we take out the first two supergames in each session, where players might still have been learning the game. This is shown in Figure 7b. Lastly, we looked at advice quality per learning level in the treatment group. The results can be seen in Figure 7c. It turns out that advice in the learning levels p_0, p_1, p_2 and p_4 is significantly higher than that given in the control group. However, the average advice quality in learning level p_3 is lower than in the control group. A potential explanation for the low advice quality in this level is the gambling effect: Experts feel that their signal strength is sufficiently high to generate fitting advice on the spot such that they will take their bonus and hope to appease the consumer in the next period. It is also noteworthy that the advice quality in learning level p_4 is significantly higher than in all other learning levels as well as in the control group. This effect cannot be explained by a better signal quality, since advice quality is measured by the share of tradeoff-situations (bonus option = option 2), in which the adviser decides to give useful advice instead of receiving his bonus. The signal quality only affects how often this decision will actually translate to the intended payoff of one to the consumer. A reason for the high advice quality in learning level p_4 could be a selection effect: The majority of advisers who reached learning level p_4 in their advice relationship probably did so because they gave good advice in the past and they might have an intrinsic motivation to give good advice and/or value long-lasting relationships a lot. Another explanation could be reciprocity: Advisers reward consumers for their loyalty over the last rounds by giving better advice.

Overall, we conclude that the difference in advice quality between control group and treatment group seems to be quite robust.

C.2. Welfare Analysis



(a) Short supergames taken out (<3 rounds)

(b) Consumers with max. 1 incorrect answer

(c) First two supergames taken out

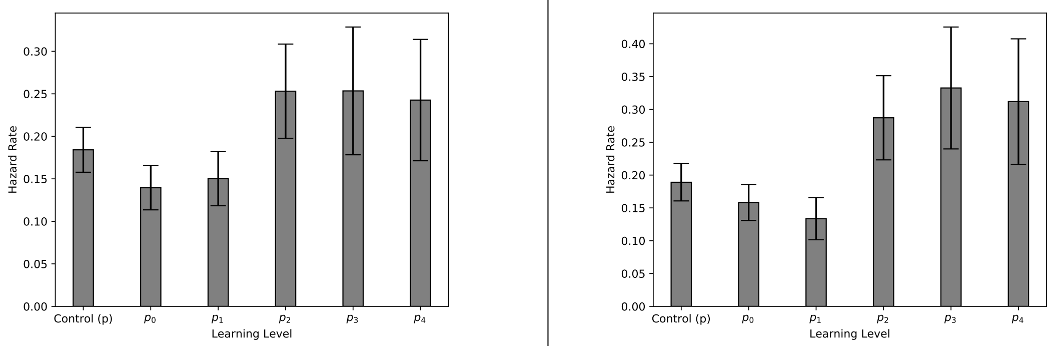
(d) Total Welfare

Figure 8: Robustness Checks for Consumer Welfare

The results of the robustness checks for consumer welfare can be seen in Figure

8. Overall, the observation that consumer welfare does not significantly differ between control and treatment group is very robust. When we take out the supergames with less than three rounds (Figure 8a) or the first two supergames of each session (Figure 8c) or those consumers with two or more incorrect answers to check questions (Figure 8b), there is no significant difference in consumer welfare between control and treatment group. We also had a look at total welfare, the sum of consumer and expert payoffs. As can be seen in Figure 8d, there is no significant difference between control and treatment group, either. This also implies that expert payoffs in control and treatment group are not significantly different from each other.

C.3. Hazard Rates



(a) Consumers with max. 1 incorrect answer

(b) First two supergames taken out

Figure 9: Robustness Checks for Hazard Rates

Figure 9 shows the robustness checks we performed for the hazard rates. Our finding that hazard rates are significantly lower in learning levels p_0 and p_1 proves to be robust. Both excluding consumers with two or more incorrect answers to check questions (Figure 9a) and taking out the first two supergames of each session (Figure 9b) leads to a shape very similar to the one in Figure 4.

References

Acquisti, A., L. Brandimarte, and G. Loewenstein (2015). Privacy and human behavior in the age of information. *Science 347*(6221), 509–514.

Benabou, R. and G. Laroque (1992). Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility. *Quarterly Journal of Economics 107*(3), 921–958.

Blume, A., E. K. Lai, and W. Lim (2020). Strategic information transmission: A survey of experiments and theoretical foundations. In *Handbook of experimental game theory*, pp. 311–347. Edward Elgar Publishing.

Byers, S., L. F. Cranor, D. Kormann, and P. McDaniel (2004). Searching for Privacy: Design and Implementation of a P3P-Enabled Search Engine. In *International Workshop on Privacy Enhancing Technologies*, pp. 314–328. Springer.

Çetin, G. S., W. Dai, Y. Doröz, W. J. Martin, and B. Sunar (2016). Blind Web Search: How far are we from a privacy preserving search engine? *IACR Cryptology ePrint Archive 2016*, 801.

Chellappa, R. K. and R. G. Sin (2005). Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. *Information technology and management 6*(2), 181–202.

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

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

Edelman, B., M. Ostrovsky, and M. Schwarz (2007). Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords. *American Economic Review 97*(1), 242–259.

Edelman, B. and M. Schwarz (2010). Optimal Auction Design and Equilibrium Selection in Sponsored Search Auctions. *American Economic Review 100*(2), 597–602.

Eliaz, K. and R. Spiegler (2011). A Simple Model of Search Engine Pricing. *Economic Journal 121*(556), 329–339.

Gramb, M. and C. Schottmüller (2022). Anonymous or personal? Repeated personalized advice in the lab. *AEA RCT Registry*. January 13. doi:10.1257/rct.8682.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association 1*(1), 114–125.

Hillenbrand, A. and S. Hippel (2019). Strategic Inattention in Product Search. *MPI Collective Goods Preprint* (2017/21).

Holmström, B. R. (1978). *On Incentives and Control in Organizations.* Stanford University.

Holmström, B. R. (1982). *On the Theory of Delegation.* Northwestern University.

Inderst, R. and M. Ottaviani (2009). Misselling through Agents. *American Economic Review 99*(3), 883–908.

Inderst, R. and M. Ottaviani (2012a). Financial Advice. *Journal of Economic Literature 50*(2), 494–512.

Inderst, R. and M. Ottaviani (2012b). How (not) to pay for advice: A framework for consumer financial protection. *Journal of Financial Economics 105*(2), 393–411.

Li, J., N. Matouschek, and M. Powell (2017). Power Dynamics in Organizations. *American Economic Journal: Microeconomics 9*(1), 217–41.

Lipnowski, E. and J. Ramos (2020). Repeated delegation. *Journal of Economic Theory 188*, 105040.

Park, I.-U. (2005). Cheap-Talk Referrals of Differentiated Experts in Repeated Relationships. *RAND Journal of Economics 36*(2), 391–411.

Schottmüller, C. (2019). Too good to be truthful: Why competent advisers are fired. *Journal of Economic Theory 181*, 333–360.

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

Sobel, J. (2013). Giving and Receiving Advice. *Advances in economics and econometrics 1*, 305–341.

Spiekermann, S., J. Grossklags, and B. Berendt (2001). E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pp. 38–47. ACM.

Tsai, J. Y., S. Egelman, L. Cranor, and A. Acquisti (2011). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study. *Information Systems Research 22*(2), 254–268.

Xu, Y., K. Wang, B. Zhang, and Z. Chen (2007). Privacy-enhancing personalized web search. In *Proceedings of the 16th international conference on World Wide Web*, pp. 591–600. ACM.

Zimmer, M. (2008). Privacy on Planet Google: Using the Theory of "Contextual Integrity" to Clarify the Privacy Threats of Google's Quest for the Perfect Search Engine. *J. Bus. & Tech. L. 3*, 109.