Experts and Their Records*

Alexander Frankel
Yahoo! Research, Stanford GSB

Michael Schwarz
Yahoo! Research

April 10, 2009

Abstract

Consider an environment where long-lived experts repeatedly interact with short-lived customers. In periods when an expert is hired, she chooses between providing a profitable major treatment or a less profitable minor treatment. The expert has private information about which treatment best serves the customer, but has no direct incentive to act in the customer's interest. Customers can observe the past record of each expert's actions, but never learn which actions would have been appropriate. We find that there exists an equilibrium in which experts always play truthfully and choose the customer's preferred treatment. The expert is rewarded for choosing the less profitable action with future business: customers return to an expert with high probability if the previous treatment was minor, and low probability if it was major.

If experts have private information regarding their own payoffs as well as what treatments are appropriate, then there is no equilibrium with truthful play in every period. But we construct equilibria where experts are truthful arbitrarily often as their discount factor converges to one.

1 Introduction

In many economic environments, uninformed customers must rely on experts to both diagnose and treat their problems. Doctors, dentists, mechanics, and management consultants all help to determine what services their clients need in addition to providing those services. There is a misalignment of incentives when experts earn higher profits on certain treatments than on others.

This paper considers a repeated environment in which experts are long-lived and customers can use experts' records of past actions to determine whom to hire. We take treatments to be pure *credence goods*: customers observe past treatments, but they never receive

*We would like to thank Heski Bar-Isaac, Martin Cripps, Drew Fudenberg, Andrzej Skrzypacz, and Steve Tadelis for their helpful comments, as well as seminar audiences at Stanford and Stanford GSB.

signals about what treatments would have been appropriate. This makes it difficult for customers to punish experts for past dishonesty, as dishonesty is never revealed. Moreover, we assume that an expert's payoffs from each action are completely independent of the customer's underlying need, and that customers are short-lived players for whom long-term contracts are impossible.

Since experts' payoffs do not depend on the problem a customer faces, an expert will play truthfully in a period only if her expected discounted profit is equal across actions. An honest expert is likely to have a balanced record over time, with the proportion of major and minor treatments close to the probability that each is needed, but it will not be an equilibrium for customers to choose the expert whose record is most balanced. If customers chose in this way, then experts would just take actions to keep their records balanced.

Experts may have the proper incentives to be truthful if customers give more business to experts who have chosen the less profitable treatments in the past. The logic is illustrated by an anecdote about McKinsey & Co.'s former managing director Marvin Bower, relayed in a Business Week obituary. "In the 1950s, Bower was summoned to Los Angeles by billionaire Howard Hughes, who wanted him to study Paramount Pictures.... But Bower sensed that nothing good could come of working for Hughes. He found the entrepreneur's approach to business 'so unorthodox and so unusual' that he felt he would never be able to help Paramount. Instead of taking the assignment and reaping a big fee, he walked away. The move was classic Bower. He built McKinsey into a global consulting powerhouse by insisting that values mattered more than money" (Byrne (2003)). In other words, by publicly rejecting a profitable action, McKinsey increased its future business.

We study a repeated game between customers and experts modeled after the above interactions and look for conditions under which the experts may be truthful in equilibrium.

In each period, a new customer arrives on the market and chooses an expert. There are two possible states of the world: the customer might need a major treatment, or a minor one. The customer prefers that the appropriate action be taken but has no information about the state of the world. His only action is to choose an expert. Once a choice is made, the customer must defer to the expert's judgment.

The chosen expert observes the state and then decides whether to provide a major or minor treatment, and the expert's payoffs depend on what she does but not on what the customer needs. In other words, her payoff is a function of the action but not the state. This can be thought of as an environment where prices are exogenously fixed at industry standard levels. Experts always prefer some work to no work, and when chosen they will earn a higher profit on major treatments than on minor ones.

While true states are never revealed to future customers, we assume the full history of experts chosen and actions taken to be observable. Experts are infinitely long lived, but customers disappear from the market after they receive a treatment.

In Section 3 we show that if experts are homogeneous then a truthful equilibrium of the repeated game can be found; the promise of future business removes the incentive to play major treatments over minor ones. Customers only need to look at the most recent action taken. If it was a minor treatment, they return to the last period's expert with high probability. If it was a major treatment, they return with a low probability. By setting appropriate probabilities, they can make experts exactly indifferent between major treatments with a high short term payoff and minor ones with a high continuation value. This explains the intuition behind the McKinsey story: it can be an equilibrium for experts to report their private information truthfully when their likelihood of future work rises with less profitable actions and falls with more profitable ones.

In Section 4 we consider a more general model with heterogeneous experts who have private information along two dimensions. In addition to observing each period's hidden state, experts privately observe their own relative payoffs from providing a minor versus major treatment. Now customers will not know what probability would make an expert indifferent across actions, and so we cannot enforce the truthful equilibria above.

Customers are short-lived and so long-term contracts are impossible. If experts could commit to long-term contracts over their actions, however, and if they did not discount future payoffs, then a quota system could allow for truthfulness in some periods. Say that contracts were written over two-period blocks, and an expert was required to play one major and one minor treatment in each block. Whatever she played in the first period, she would play the opposite in the second period and get one major payoff and one minor payoff over the block. Regardless of her relative payoffs across actions, then, the expert would agree to be truthful in the first period of any block. In a similar contract with blocks of length $K > 2$, the expert would be truthful until one of the actions reached its quota and would then play the other action deterministically to the end of the block. If the quota for action a were set to be close to K times the probability of a being appropriate, then the share of truthful actions would approach 100% as K grew large.

Even if long-term contracts were possible, the discounting of payoffs would prevent us from exploiting this idea directly. Under a quota like the one described above, experts would shift all of the profitable major treatments into the early periods. One way to return experts to truthful play in early periods would be to allow their payoffs to depend slightly on the state of the world, and in particular to suppose that experts receive a utility benefit from taking the appropriate action in a period. Consider an expert who observes that a minor treatment is appropriate in an early period of a block. Playing the minor action over the major one would impose a cost from delaying the major payoff, but would yield a benefit from aligning the action with the state in this period. If she were sufficiently patient then the cost would be small, so she would prefer to be truthful and would play the minor action.

In Section 4 we show that the logic of a quota can be recovered to induce heterogenous experts to act truthfully in almost every period, even in an environment in which there is no commitment over time and experts have no preference for aligning actions with states. In a standard quota, the number of plays on an action is constant over prospective equilibrium paths; once one action hits its quota, the underplayed action must be played until the end of the block. Instead of a literal quota where the number of times an action is played is constant over prospective equilibrium paths, we use a "discounted quota" in which the number of *expected discounted* plays is constant over paths. In each block, experts are truthful in early periods, and in later periods they deterministically choose the underplayed action until a new block begins.

Here is a basic example to illustrate the idea of such a discounted quota (see Example 1). Strategies repeat every three periods. In the first period the expert acts truthfully, performing a major or minor treatment. In the second and third periods any expert who is chosen plays the opposite action of the first period, regardless of the state. After the first (truthful) period, the new customer keeps the old expert with some probability q and moves to a new one with probability $1-q$. After the second and third periods, the customer retains the expert if the suggested action was played and otherwise goes to a new one.

Say that experts have discount factor β in this example. Over the course of the three periods, an expert chosen in the first period gets an expected discounted weight of 1 play towards whichever action is taken first, and a weight of $q(\beta+\beta^2)$ towards the opposite action. If the retention probability q is $\frac{1}{\beta+\beta^2}$ then the weight is 1 on both actions. So over the three periods an expert gets one major payoff and one minor payoff in expectation along either path, and is willing to condition her first period action on the state of the world.

The customer facing a truthful expert is happy, and has no incentive to deviate. But customers facing deterministic experts are stuck – whenever a customer switches experts, the new expert plays exactly like the old one would have.

The example allows for truthfulness in every third period. Taking the experts' discount factor to 1 and enforcing the discounted quotas over longer and longer blocks, we can construct equilibria with an arbitrarily high share of truthful periods.

In this paper, we examine whether experts can be induced to act truthfully under very strict assumptions: experts have no intrinsic motivation to aid customers, and customers never learn whether past experts had been truthful. Even after stacking the deck in this manner, we find that truthful play is still possible so long as records are available. In more realistic settings experts may be somewhat altruistic, or customers (one-time or repeat visitors) may observe signals about the quality of past play. In either of these cases, our results provide a lower bound on what is achievable; the equilibria we construct continue to hold.

2 Literature Review

Darby and Karni (1973) introduced the concept of credence goods, goods whose value is known by a seller but never fully revealed to the consumer. Dulleck and Kerschbamer (2006) provide a recent review of the literature on when and how credence goods can be provided efficiently. In the credence goods literature, most work focusses on inducing truthfulness in one-shot settings in which expert payoff levels are common knowledge.

In the terminology of Dulleck and Kerschbamer (2006), we impose the Verifiability rather than Liability assumption: customers can confirm that the announced treatment has been performed, but the success or failure of the treatment is not publicly observed and is noncontractible. When the Liability assumption holds instead, experts will correctly treat the problems but may attempt to overcharge customers, performing a cheap treatment but reporting an expensive one.

We also impose what that paper calls the Commitment assumption, that a customer who goes to an expert must be treated by that expert. When this is relaxed, truthfulness can be induced by having one expert diagnose the problem and another perform the treatment. This is explored in Wolinsky (1993), Dulleck and Kerschbamer (2006, 2008), and Alger and Salanie (2006). In these models, agents may incur inefficient search and diagnosis costs. An alternative way of relaxing the Commitment assumption is to prevent the customer from seeking other experts, but allowing him to refuse treatment after observing a diagnosis. Pitchik and Schotter (1987) take this approach and find a mixed strategy equilibrium in which customers sometime reject expensive treatments and experts are sometimes truthful.

If the prices for treatments are set so that profits are equal across all actions, then experts will play truthfully. Emons (1997) and Dulleck and Kerschbamer (2006) build models of credence goods which exploit this solution. In this sense, when the parties can bargain over prices they may achieve truthful play without resorting to a repeated game. But such an approach only works when, as in our model of homogeneous experts in Section 3, experts have no private information about their costs for each action. If costs are privately observed, as in Section 4, then customers have no way to know what prices would induce truthful play, and experts will have no incentive to report their costs honestly. In the Web Appendix, we elaborate on the impossibility of endogenous prices to induce truthful play in a one-shot setting.

There is a large body of work on repeated games outside of the context of credence goods. It is common for players to have private information on their own payoffs, as experts do in Section 4. Our model diverges from a standard set-up in that some players – the customers – do not know their own payoffs over others' actions.

Bar-Isaac and Tadelis (2008) provide a recent survey of results on repeated games with "reputations" in which some players have hidden types and other players have beliefs about these types. The uninformed agents observe past actions (or signals thereof) to generate inferences about others' types, and therefore about the future actions they may play. When we consider the case where experts have private information on their own payoffs in Section 4, we will construct equilibria where reputational dynamics are trivial: all experts play identical strategies, and customers make no inferences about types or strategies based on histories.

Bar-Isaac and Tadelis (2008) also discuss the problem of firms' trying to develop a reputation for having expertise, which – although orthogonal to our analysis – is more in the spirit of credence goods. To demonstrate expertise, firms may may have incentives to skew their reports to match or go against what customers expect to hear.

There is a small literature involving repeated markets for credence goods.

Fudenberg and Levine (1994) present a number of general results about payoff frontiers in repeated games with long-run and short-run players who do not have private payoff types. Their Investment Game example shares a number of features with our model of homogeneous experts in Section 3. As in our paper, short-lived players offer business to long-lived players who may secretly take advantage of them. In equilibrium, the long-run players' temptation to cheat can largely be overcome by the threat that future short-run players will withdraw their business after suspect outcomes. Our model allows us to explicitly construct efficient equilibria; Fudenberg and Levine (1994) focus on conditions under which there exist equilibria approaching efficiency.

Schneider (2007) studies repeated interactions in the market for car repair, an example of a credence good. He considers a 2-period model with multiple experts, and runs a field experiment to test the predictions. As in our paper, Schneider (2007) takes prices to be fixed exogenously and shows that there is an equilibrium in which customers return to an expert with lower probability after an expensive repair. In this equilibrium, profit-maximizing mechanics are honest in the first period and do unnecessary major repairs in the second period. Our equilibrium in Section 3 demonstrates a similar intuition for inducing truthfulness in every period in a fully repeated setting.

Wolinsky (1993) also considers a setting where customers return to a market twice and the choice of expert in the second period depends on the expert's first period action. In this model experts may reject customers with expensive problems, and customers return to experts who had been willing to treat them in the past.

Park (2005) studies an infinitely repeated game in which there are finitely many expert mechanics, and a diagnosis reveals to the mechanics which of them can best perform the repair. After a diagnosis, mechanics report their diagnoses (possibly falsely) and the customer chooses an expert based on the reports. Experts have no opportunities to lie about the actual repair they perform; once an expert is chosen, she fixes the car and receives a deterministic payoff. Park (2005) finds that equilibria with many or even 100% truthful reports are possible when payoffs do not vary too much across periods. One crucial feature of the model is that customers learn the true state at the end of a period, which lets them punish liars. This means that experts do not have to be made precisely indifferent over reports in order to be truthful, and so the equilibria are robust against some uncertainty in the experts' payoffs.

Ely and Valimaki (2003) study a model where short-lived customers play a repeated game with long-lived mechanics who privately observe the state and determine the proper repair. In their model altruistic mechanics strictly prefer to act truthfully and perform repairs which are appropriate to the state, while bad mechanics always want to do engine repairs. But instead of being truthful at first, the good mechanics will do a tune-up in order to separate themselves from bad mechanics and prove their goodness to all future customers. No consumer wants to be the first to go to an expert, so the market breaks down. Ely, Fudenberg, and Levine (2008) extend this work and find sufficient conditions for when observable histories lead to market collapse.

While we set up the problem similarly to Ely and Valimaki (2003), two key differences make our strategic environment vastly different. First, there is no altruism in our model – the utility an expert receives from each repair is independent of the state of the world.

Second, in their model customers will exercise an outside option rather than receive a repair that is independent of the state; we do not allow customers to opt out of the market. This prevents the market from breaking down.

Ely and Valimaki (2003) and Schneider (2007) consider the possibility that some proportion of experts are altruistic. One of the key innovations of our paper is to treat a different form of unobserved heterogeneity, in which experts are profit maximizers but are privately informed about their payoffs across actions. Park (2005) does allow for private information on instantaneous profits, but in that paper customers can discover and punish deviations. The punishments give experts a strict incentive to prefer truthful reports over a range of profit levels.

Finally, there is another set of related papers which bears mentioning. Recall that in our model an expert can never be given a strict incentive to act truthfully. We can achieve truthful play via indifference by enforcing "discounted quotas" which fix the number of expected discounted times that each action can be played by an expert. Past work has explored the use of standard quotas, which fix the absolute number of times that an action is played, to induce truthful revelation of private information. Townsend (1982) shows how quotas can be applied to the context of repeated bilateral trade, and Jackson and Sonnenschein (2007) extends this to a general environment with independent and ex ante identical allocation decisions. Agents are asked to report types jointly over many decisions, and the distribution of reported types is restricted to match the theoretical distribution. Jackson and Sonnenschein (2007) call this the "linking" of separate decisions through "budgets" or "rations." As more decisions are linked, the mechanisms approach efficiency.

As mentioned in the introduction, when agents discount future payoffs a standard quota will only work if agents receive some benefit from telling the truth. In these papers there is such a benefit because the agents' private information regards their own preferences, and conditional on their reported types efficient outcomes are realized. If a trade is more likely when a buyer reports a high value, and if a buyer can only report that he has a high value a limited number of times, he prefers to report this when his value truly is high. In both papers, the mechanisms fall apart if agents cannot commit in advance to participate over long time horizons.

Although less directly related to our own work, Pesendorfer (2000) employs a similar intuition. He studies a bidding cartel for procurement contracts in which firms link separate auctions by reporting to each other a desirability ranking of the available contracts, and determine bidders from the cartel based on these rankings. As the number of linked auctions increases, the mechanism gives rise to nearly optimal bids. A quota is analogous to a ranking when information is binary rather than continuous.

3 Homogeneous Experts

There is a set of experts $E = \{e^1, e^2, ...\}$ and a set of customers $C = \{c^1, c^2, ...\}$. Customers are short-term players, while experts are infinitely long-lived and have discount factor $\beta \in (0, 1)$.[1]

In period $t \in \{1, 2, ...\}$, customer c^t arrives on the market and observes the past history of experts chosen and actions taken. The customer then chooses a single expert e_t from E. (Superscripts denote elements of the set E, while subscripts represent time periods). The expert observes the state θ_t and then chooses an action a_t.

In each period the set of possible states is $\Theta = \{\theta^m, \theta^M\}$, and the set of actions for the chosen expert is $A = \{m, M\}$, where m refers to a "minor" treatment and M a "major" one. The customers always want the expert to be "truthful" and choose a when the state is θ^a. But in the short term, every expert prefers action M. Formally, write stage payoffs in period t as

$$\begin{aligned} \text{Customer } c^t: &\quad U^t(a_t|\theta_t) \\ \text{Expert } e^i: &\quad \begin{cases} 0 & \text{if } e_t \neq e^i \\ R(m) = r & \text{if } e_t = e^i \text{ \& } a_t = m \\ R(M) = 1 & \text{if } e_t = e^i \text{ \& } a_t = M \end{cases} \end{aligned}$$

where these stage payoffs satisfy

$$\begin{aligned} &U^t(a|\theta^a) > U^t(a'|\theta^a) \text{ for } a' \neq a \\ &0 < r < 1. \end{aligned}$$

For each expert only the relative payoffs of the different actions matter, so we have normalized $R(M)$ to 1 and the payoff when not selected to 0. $R(m) = r$ is in between these two. Expert e^i's lifetime utility is $\sum_{\{t|e_t=e^i\}} \beta^{t-1} R(a_t)$.

A customer only receives a payoff in the period in which she chooses an expert, and this payoff is a function of the treatment received along with the current state of the world. The

[1] The set of experts is modeled as countably finite, but our arguments will not hinge on this assumption; we discuss the case of finitely many experts in the Web Appendix.

payoff does not depend on the identity of the expert.

Although we do not analyze efficiency explicitly, efficiency is synonymous with truthfulness if the benefit to customers from an appropriate treatment always outweighs the costs to the experts.

In each period, the state is θ^m with probability $0 < p < 1$ and θ^M with probability $1 - p$.

The identities of experts chosen in past periods and the actions taken by these experts are publicly observable to all players. We write this list of experts and actions observed prior to period t as a "public history" $H_t = (e_1, a_1, ...e_{t-1}, a_{t-1})$, with $H_1 \equiv (\emptyset)$. Let $\mathcal{H}_t$ be the set of all possible public histories at time t, and let $\mathcal{H}$ be the set of possible public histories at any time: $\mathcal{H} \equiv \cup_t \mathcal{H}_t$.

Let $\oplus$ be the concatenation operator, so that for histories H and H', the notation $H \oplus H'$ means history H followed by history H'. Say that a history G begins with history H if $G = H \oplus H'$ for some H'.

Customers observe the list of past experts and actions, but they have no way of discerning whether past actions were appropriate. A customer's only decision is to use the observable public history to choose an expert to treat his problem. The customer may play a mixed strategy and choose experts probabilistically. We write customer c^t's strategy in period t as

$$\rho^t : \mathcal{H}_t \to \Delta(E)$$

where, for any countable set S, $\Delta(S)$ denotes the space of probability distributions over S. In order to avoid awkward descriptions of pure strategies, we will slightly abuse notation and use s to denote the element of $\Delta(S)$ which places probability 1 on $s \in S$. Define the collective strategies of all customers as $\rho \equiv (\rho^1, \rho^2, ...)$.

Each expert also sees the public history, and once chosen, she also observes the current state. The expert then chooses a treatment based on all of this information. We write expert e^i's strategy conditional on being chosen as

$$\sigma^i : \mathcal{H} \times \Theta \to \Delta(A).$$

More generally, an expert could also condition her strategy on privately observed values of the state in previous periods in which she was chosen. Past states are payoff irrelevant to all players, so allowing this would complicate notation without affecting our results.[2]

[2] Any strategy that is optimal in the class of those which do not depend on past states is also optimal in the larger class of strategies which do. So the equilibria we construct will remain equilibria in the more general

In this model, the expert's utility is independent of the true state, and the customers cannot confirm whether an expert has acted honestly or dishonestly in the past. Moreover, each customer is a short term player who is unable to reward or punish an expert after choosing her.

There exists an inefficient equilibrium in which every expert always performs the more profitable major treatment M. We will show that when experts are patient enough a truthful equilibrium will also exist, in which experts always play the action corresponding to the true state. Here and in the rest of the paper, the term equilibrium refers to a sequential equilibrium.

Definition. The expert e^i with strategy σ^i is said to be *truthful* at history H_t if $\sigma^i(H_t, \theta^a) = a$ for each $a \in A$. A *truthful equilibrium* is an equilibrium in which, at every equilibrium history, every expert who may be chosen with positive probability is truthful.

Proposition 1. *A truthful equilibrium exists if and only if* $\beta \geq 1 - r$.

Before proving the proposition, it is useful to state a lemma. Recall that an expert's payoffs are independent of the state, so an expert is only willing to play truthfully if she is indifferent across possible actions. The following lemma states that if truthful play on some set of histories is a best response, then the strategy remains a best response if the expert switches to arbirtrary play at those histories.

Lemma 1. Fix the customers' strategy ρ and the strategies σ^{-i} of experts aside from e^i, and suppose that expert e^i has a sequentially rational best response σ^i. Take $\hat{\mathcal{H}} \subseteq \mathcal{H}$ to be a set of histories for which σ^i is truthful. Now construct a new strategy $\hat{\sigma}^i$, where $\hat{\sigma}^i : \mathcal{H} \times \theta \to \Delta(A)$ is identical to σ^i on all $\mathcal{H} \backslash \hat{\mathcal{H}}$, and is arbitrary on $\hat{\mathcal{H}}$. The strategy $\hat{\sigma}^i$ gives the same expected utility as σ^i starting at every history and is also a sequentially rational best response.

Proof. See Appendix. □

Proof of Proposition 1. First, we will show the "If" direction. Suppose that $\beta \geq 1 - r$; we will now construct a truthful equilibrium.

Experts play truthfully: for each $a \in A$, for each $e^i \in E$, and for each $H_t \in \mathcal{H}$, let $\sigma^i(H_t, \theta^a) = a$.

strategy space. The results that no truthful equilibria can exist under various conditions (see Proposition 1 and Remark 1) also hold in the general strategy space.

The customer's strategy is the following. c^1 chooses $\rho^1(H_1) = e^1$. For $t > 1$, c^t chooses

$$\rho^t\big(H_{t-1} \oplus (e_{t-1} = e^i, a_{t-1})\big) = \begin{cases} e^i \text{ with Prob } q(a_{t-1}) \\ e^{i+1} \text{ with Prob } 1 - q(a_{t-1}) \end{cases}$$

with $q(m) = 1, q(M) = \frac{r+\beta-1}{r\beta}$. Notice that $0 \leq q(M) < q(m) = 1$.

Because all experts play truthfully at every history, customers are indifferent across experts and any strategy is a best response.

To check that truthful play is a best response for the experts, consider the expected payoff V for an expert who follows the strategy, conditional on being chosen in a given period but unconditional on the realization of θ. By the one-shot deviation principle, the expert's strategy is optimal if she always plays a maximizer of $R(a) + \beta q(a)V$, and if $V = \max_{a \in A}\{R(a) + \beta q(a)V\}$. This holds if we can find a V such that

$$\begin{aligned} & V = 1 + \beta \frac{r+\beta-1}{r\beta} V = r + \beta V \\ \iff & V\left(1 - \frac{r+\beta-1}{r}\right) = 1 \;\&\; V(1-\beta) = r \\ \iff & V = \frac{r}{1-\beta}. \end{aligned}$$

So the continuation value V is $\frac{r}{1-\beta}$, and there are no profitable deviations.

This shows that the above strategies are an equilibrium when $\beta \geq 1 - r$ – the experts and customers are indifferent with respect to all actions at all histories.

Now, in the "Only If" direction, suppose that a truthful equilibrium exists. By Lemma 1, if a given expert is willing to play truthfully at every period in which she is chosen then she must be indifferent to switching to the strategy of playing m at every period, or to the strategy of playing M in every period. An expert selected in the current period whose strategy is to always play m will get stage payoffs of at most r in each period (0 in any period in which she is not chosen) and so her present value of future payoffs is at most $\frac{r}{1-\beta}$. An expert who plans to play M in every period gets a stage payoff of 1 today, and some nonnegative payoff in the future. The expert can only be indifferent over these two strategies if $\frac{r}{1-\beta} \geq 1$, or rather $\beta \geq 1 - r$. $\square$

To implement this equilibrium customers only need to observe the previous period's action, and experts can ignore the histories entirely. The customer utility functions never appear in the construction. Moreover, the equilibrium is completely independent of p, the parameter controlling the probabilities of the various states of the world. The expert's continuation probabilities from different actions are set such that every expert is exactly indifferent between the higher payoff today versus the lower payoff in the future from playing M, and p plays no role in this.

In the above equilibrium, customers return to an expert with probability 1 if the minor action m was played last period, and probability less than 1 if the major action M was played. In the Web Appendix we show that similar truthful equilibria can be constructed if expert utility functions differ but are observed by customers, or if experts have more than two possible actions.

4 Heterogeneous Experts

In Section 3 customers always return to an expert who has just performed a minor treatment, and rehire an expert who has just performed a major treatment with a probability less than 1. This probability is chosen so that each expert will be exactly indifferent between a major and minor treatment in every period. But this equilibrium falls apart if customers are no longer certain about an expert's relative payoff from the two treatments. An expert with a slightly lower payoff from the major treatment will find it profitable to play the minor treatment in each period she is chosen, and vice versa.

We will proceed to consider whether any truthful play is possible when an expert's payoffs are private information. We find that experts can often be incentivized to play truthfully when strategies depend on more than just the last period's play.

In this section, we take expert stage payoffs to be

$$\text{Expert } e^i : \begin{cases} 0 & \text{if } e_t \neq e^i \\ R^i(m) = r^i & \text{if } e_t = e^i \ \& \ a_t = m \\ R^i(M) = 1 & \text{if } e_t = e^i \ \& \ a_t = M \end{cases} .$$

We maintain the normalization of the instantaneous payoff of the major treatment to 1, and we now let the relative payoff of the minor to major treatment be an expert's private information. Each expert e^i realizes a relative payoff r^i drawn from a distribution over $\mathbb{R}_+$, the set of nonnegative real numbers; some (or all) experts may prefer m to M. These distributions need not be identical or independent. The realization of r^i is privately observed by expert e^i at the start of the game, and is fixed over time.

Experts have discount factor β as before, and β is common knowledge.

Customer payoffs are as in Section 3. Customers prefer to receive appropriate treatments, but conditional on the action and state have no preferences over the identity of the expert. Each customer still observes only the public history of past experts and actions when choosing a new expert, so customer strategies are also as above; the collective strategy of the customers is $\rho : \mathcal{H} \to \Delta(E)$.

We now have to generalize the strategy space of each expert to depend not only on the history and state but also her realized type r^i. In this context, an expert e^i's strategy is a function $\phi^i : \mathcal{H} \times \Theta \times \mathbb{R}_+ \to \Delta(A)$. For ease of notation, we will use the term *conditional strategy* to refer to maps $\sigma : \mathcal{H} \times \Theta \to \Delta(A)$. An expert's strategy can be thought of as a map from types to conditional strategies. The conditional strategy for an expert e^i of type $r \in \mathbb{R}_+$ will be denoted by $\sigma^i_r(H^i, \theta) \equiv \phi^i(H_t, \theta, r)$.

At any history H_t with state θ_t at which the expert $e_t = e^i$ is chosen, a conditional strategy σ for the expert induces *weights* on the number of expected discounted times that m and M will be played in the current and future periods. For $a \in A$, we can define the weight on action a by

$$W^i_a(H_t, \sigma | \theta_t, \phi^{-i}, \rho) \equiv \sum_{\tau=t}^{\infty} \beta^{\tau-t} \text{Prob}[e_\tau = e^i \text{ and } a_\tau = a | \theta_t, \sigma, H_t, \phi^{-i}, \rho; e_t = e^i]$$

where the probability is taken with respect to the (Bayesian) beliefs of e^i . The expected present value of conditional strategy σ at some history H_t (conditional on $\phi^{-i}, \rho, \theta_t$, and conditional on $e_t = e^i$, but suppressing these from the notation) in the current and future periods is

$$r^i \cdot W^i_m(H_t, \sigma) + W^i_M(H_t, \sigma).$$

Sequential rationality requires that $\sigma^i_{r^i}$ be a maximizer of this expression for each $r^i \in \mathbb{R}_+$. Notice that this value depends on θ_τ in periods where $e_\tau = e^i$ only through the effect of θ_τ on σ.

It can be natural to think of an expert e^i as choosing a bundle (W^i_m, W^i_M) at each action node rather than a strategy. Let the set of available bundles be denoted

$$\mathcal{W}^i(H_t) \equiv \{(W^i_m(H_t, \sigma), W^i_M(H_t, \sigma)) | \sigma : \mathcal{H} \times \Theta \to \Delta(A)\},$$

suppressing the dependence on ϕ^{-i} and ρ. $\mathcal{W}^i(H_t)$ is independent of θ_t because no matter what the state is, a strategy exists that plays "as if" the state were the opposite.

Points in $\mathcal{W}^i(H_t)$ fully determine an expert's utility going forward, so we can consider indirect preferences over this set. Expert e^i's indifference curves over $\mathcal{W}^i(H_t)$ are straight lines with slope $-r^i$. The equilibrium in Proposition 1 holds up because the customers construct $\mathcal{W}^i(H_t)$ such that all points lie on a straight line with slope $-r$; every expert is indifferent between every action at every history.

No experts with different values of r^i can be indifferent over distinct pairs of (W^i_m, W^i_M) because the indifference curves have a unique intersection point. However, if two distinct strategies yield the same pair of weights, then experts of any $r^i \in \mathbb{R}_+$ will be indifferent. We formalize this in the following lemma.

Lemma 2. Take two conditional strategies $\sigma', \sigma'' : \mathcal{H} \times \Theta \to \Delta(A)$, and take $r' \neq r'' \in \mathbb{R}_+$. An expert e^i chosen at history H_t is indifferent between σ' and σ'' for both possible types r' and r'' if and only if

$$W^i_m(H_t, \sigma') = W^i_m(H_t, \sigma'') \ \& \ W^i_M(H_t, \sigma') = W^i_M(H_t, \sigma'').$$

Proof. See Appendix. □

Lemma 1 of Section 3 continues to hold: an expert whose best response includes truthful actions is indifferent to switching her strategy arbitrarily at truthful histories.

Remark 1. If each expert can realize at least two possible types then no truthful equilibrium exists. More formally, there is no truthful equilibrium if for no expert e^i does there exist $r \in \mathbb{R}_+$ such that $r^i = r$ with probability 1.

Proof. Suppose there is a truthful equilibrium. Take e^i to be some expert who is selected with positive probability at H_1. In a truthful equilibrium, there is a probability one that r^i is such that $\sigma^i_{r^i}$ is truthful at every equilibrium history in which e^i is chosen with positive probability. Take some such r^i. (An equilibrium history is an element of $\mathcal{H}$ which occurs with positive probability in equilibrium).

For each $a \in A$, consider the conditional strategy σ_a of playing action a at every history. The payoff to the expert of playing σ_a is exactly the same as the payoff of a strategy which plays a at all equilibrium histories in which she is selected with positive probability, and mimics $\sigma^i_{r^i}$ at other histories. And by Lemma 1, such a strategy is optimal; the expert is truthful at all equilibrium histories in which she may be chosen, and so is indifferent to modifying her strategy arbitrarily at these periods. So the expert e^i of type r^i is indifferent between σ_m and σ_M at H_1.

But the two conditional strategies σ_m and σ_M induce different points in $\mathcal{W}^i(H_t)$: if expert e^i is chosen at H_1, then $W^i_M(H_1, \sigma_M) \geq 1$ while $W^i_M(H_1, \sigma_m) = 0$. So by Lemma 2, there is at most a single type in $\mathbb{R}_+$ for which e^i would be indifferent between σ_m and σ_M at H_1. This type must be realized almost surely in a truthful equilibrium, and so there is no truthful equilibrium if no single type is realized with probability 1. □

While an equilibrium with truthful play at every history cannot exist, we can still find equilibria in which experts play truthfully at some histories. Here is an example with truthful play once every third period.

Example 1. We will show that for any $\beta \gtrsim .80$, the following strategy profile constitutes an equilibrium in which experts are truthful at periods $t = 1, 4, 7, 10$, and so on.

First we define T_0 be the set of time periods of the form $3n + 1$, and $t_0(t) \in T_0$ to be the most recent period in T_0 up to period t:

$$\begin{aligned} T_0 &\equiv \{\tau | \tau = 3n + 1 \text{ for some } n \geq 0\} \\ t_0(t) &\equiv \max\{\tau \in T_0 | \tau \leq t\}. \end{aligned}$$

So $t_0(t)$ is 1 for $t = 1, 2, 3$, and $t_0(t) = 4$ for $t = 4, 5, 6$. It will be convenient to use t_0 for both the function $t_0(t)$ and as a representative element of the set T_0.

The experts' strategies are as follows. All experts of all types play identical conditional strategies. They will be truthful at the T_0 periods, experts will play the action opposite of that played at $t_0 \in T_0$ in the periods $t_0 + 1$ and $t_0 + 2$. Strategies repeat every three periods. So for all $e^i \in E$, for all $r \in \mathbb{R}_+$, and for each $a \in A$,

$$\sigma^i_r(H_t, \theta^a) = \begin{cases} a & \text{if } t_0(t) = t \\ M & \text{if } t_0(t) > t \ \& \ a_{t_0(t)} = m \\ m & \text{if } t_0(t) > t \ \& \ a_{t_0(t)} = M \end{cases}.$$

See Figure 1 for an illustration of the experts' strategies.

We move now to the customers' strategies. After each period, the customer either returns to the previous expert or "fires" her and moves to an entirely new one. After the first period in each repeating block, the truthful period t_0, the customer arriving at $t_0 + 1$ has some fixed positive probability of firing the old expert. After the two deterministic periods, the next customer returns to the expert if she played the suggested deterministic action and fires her otherwise. Formally, at period 1 c^1 chooses $\rho^1(H_1) = e^1$. At period $t > 1$, if the previous expert chosen was e^i (so that $H_t = H_{t-1} \oplus (e^i, a_{t-1})$), the customer chooses

$$\rho^t(H_t) = \begin{cases} e^i & \text{with Prob } q(H_t) \\ e^{i+1} & \text{with Prob } 1 - q(H_t) \end{cases}$$

where the function $q : \mathcal{H} \setminus \mathcal{H}_1 \to [0,1]$ determines the probability of continuing with an expert rather than moving to the next one. $q(H_t)$ satisfies

$$q(H_t) = \begin{cases} \frac{1}{\beta+\beta^2} & \text{if } t_0(t) = t-1 \\ 1 & \text{if } t_0(t) \neq t-1 \ \& \ a_{t_0(t-1)} \neq a_{t-1} \\ 0 & \text{if } t_0(t) \neq t-1 \ \& \ a_{t_0(t-1)} = a_{t-1} \end{cases}.$$

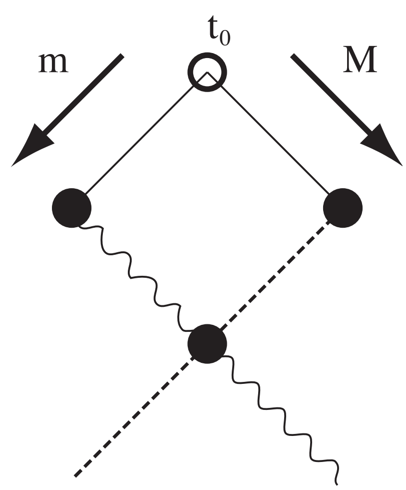

Figure 1: Expert Strategies in Example 1.

This picture illustrates the equilibrium paths of play in a single 3-period block of the strategies in Example 1. At the end of the block, the strategy repeats. The open circle represents a truthful period; the closed circles represent deterministic periods. The style of the lines is varied in order to show which histories lead to which actions at deterministic periods.

Such a strategy profile can be constructed so long as $\frac{1}{\beta+\beta^2} \leq 1$, that is, $\beta \geq \frac{\sqrt{5}-1}{2} \cong .62$. The customers are necessarily best responding because at every history, each expert plays identically to every other expert. So the customer is indifferent over who he chooses. We will show that the experts' strategies also constitute a best response when their discount factor is large enough, and therefore that this strategy profile is indeed an equilibrium.

Consider the contribution to $W^i_M(H_{t_0}, \sigma^i_r)$ and $W^i_m(H_{t_0}, \sigma^i_r)$ from the three periods $t_0, t_0 + 1, t_0 + 2$, for $t_0 \in T_0$. If $\theta_{t_0} = m$ then the expert is instructed to play m in t_0, and M in $t_0 + 1$ and $t_0 + 2$ if chosen. This gives a weight of 1 towards the minor action m and a weight of $\frac{1}{\beta+\beta^2}(\beta + \beta^2) = 1$ towards the major action M over the block. The same holds if $\theta_{t_0} = M$ and the expert is instructed to play M, m, m; the expert gets one minor payoff and one major payoff in expected discounted terms over the three periods. She is therefore indifferent between m and M at the t_0 periods and is willing to follow the suggested strategy of truthful play.

We still have to show that there are no profitable deviations at the $t_0 + 1$ or $t_0 + 2$ stages. At a period with t equal to $t_0 + 1$ or $t_0 + 2$, suppose that $\underline{a}$ is the suggested action and the selected expert considers a deviation to $\overline{a} \neq \underline{a}$. Deviating gives a weight of 1 towards $\overline{a}$ and 0 towards $\underline{a}$. So a sufficient condition for deviations to be unprofitable is that the weight on $\overline{a}$ from following the equilibrium is at least 1.

An expert chosen at t_0 receives a weight of 1 towards both m and M over the three periods of the block. Iterating this out, the weight over all periods from following the strategy is $W^i_a(H_{t_0}, \sigma^i_r) = \frac{1}{1 - \frac{1}{\beta+\beta^2}\beta^3} = \frac{1+\beta}{1+\beta-\beta^2}$ for $a = m, M$. So following the equilibrium starting at $t_0 + 1$ gives a weight of $\beta^2 \frac{1+\beta}{1+\beta-\beta^2}$ towards $\overline{a}$; starting at $t_0 + 2$, the weight is $\beta \frac{1+\beta}{1+\beta-\beta^2}$. Deviating is unprofitable as long as $\beta^2 \frac{1+\beta}{1+\beta-\beta^2} \geq 1$, or rather $\beta^3 + 2\beta^2 - \beta - 1 \geq 0$.[3]

This gives us a condition under which the proposed strategy will be a best response at all histories for all experts, and we already determined that the customers are best responding. The proposed strategy is an equilibrium as long as $\beta^3 + 2\beta^2 - \beta - 1 \geq 0$, which holds for $\beta \gtrsim .80$. ■

In the above example, we have blocks of length three in which there is a truthful period followed by two deterministic periods in each block. The weight on each action is constant across all prospective equilibrium action paths. At truthful periods different equilibrium paths allow for different actions in the current period, and so the expert is willing to condition her choice of path on the payoff-irrelevant state of the world. At deterministic periods, the continuation payoff from following the suggested strategy and receiving future work is greater than the benefit from deviating and never being chosen again.

[3] This is a sufficient but not necessary condition. We may have an equilibrium for values of β which don't satisfy this, if the support of possible values of r^i is limited to some subset of $\mathbb{R}_+$ bounded away from 0 and infinity. In particular, for any $\beta \geq \frac{\sqrt{5}-1}{2}$ there are no profitable deviations if all values of r^i are known to be in the interval $[1 - \beta, \frac{1}{1-\beta}]$.

All action paths consistent with equilibrium play give an expert the same weights towards each action; for each action, she faces a "quota" or a "budget" on the number of expected discounted plays. When the discount factor is large enough any off-equilibrium strategy provides a weakly lower weight on both actions, and hence has a weakly lower payoff for an expert of any type.

As the discount factor increases, we can find similar equilibria in which truthful periods occur more frequently. We take strategies that repeat in blocks of longer than three periods, and have experts play truthfully until some number of either m or M actions are played within the current block. Once one of these actions reaches enough plays, the opposite action is played until the end of the block. Taking the length of blocks to be large and taking the discount factor to 1, we can get the long-term proportion of truthful periods to approach 1.

Proposition 2. *Take $\epsilon > 0$. For β large enough, there is an equilibrium in which the long-run proportion of truthful periods is greater than $1 - \epsilon$ with probability 1.*

Proof. We will use the notation $\lfloor x \rfloor$ to denote the "floor" of a number x, the greatest integer less than or equal to x. First, a technical lemma:

Lemma 3. For any $k \in \mathbb{N}$, there exists $K \geq k$ such that

i. $1 < pK < K - 1$,

ii. pK is not an integer,

iii. $2p - 1 + \zeta < pK - \lfloor pK \rfloor < 2p - \zeta$ for $\zeta \equiv \min\{p/2, (1-p)/2\} > 0$, and

iv. $\min\left\{\frac{(1-p)\lfloor pK \rfloor}{p\lfloor(1-p)K\rfloor}, \frac{p\lfloor(1-p)K\rfloor}{(1-p)\lfloor pK \rfloor}\right\} > 1 - \frac{1}{4}\min\left\{\frac{p}{1-p}, \frac{1-p}{p}\right\}$.

Proof. See Appendix. □

Take some K satisfying conditions (i)-(iv) of the above lemma; we will construct a strategy profile for which strategies repeat every K periods. Conditions (i) and (ii) are necessary for constructing the strategies of the experts. Conditions (iii) and (iv) will guarantee that the proposed probabilities chosen by the customers are valid for a large enough discount factor, and also that the experts' responses are optimal. We will show that the strategy is well-defined and is an equilibrium for β large enough, and that as K is taken to ∞ the long-term proportion of truthful periods will converge to 1. Because we can find K arbitrarily large that satisfies the above conditions, this means that we can find equilibria in which truthful periods occur arbitrarily often.

Strategies will be defined on blocks of K periods, and will reset at periods of the form $nK + 1$. T_0 will denote the set of these periods at which new blocks begin, and $t_0(t)$ will be the most recent period in T_0 starting at period t:

$$T_0 \equiv \{\tau | \tau = nK + 1 \text{ for some } n \geq 0\}$$
$$t_0(t) \equiv \max\{\tau \in T_0 | \tau \leq t\}.$$

The term t_0 will express this function as well as a representative element of the set T_0.

Each block begins with a segment of truthful periods. Once m or M is played a certain number of times within the block, the players move into a segment where experts take deterministic actions. After K periods in the block, $t_0(t)$ increments up by K and strategies repeat. We now construct these strategies.

Partition $\mathcal{H}$ into "deterministic histories" $\mathcal{H}(D)$ and "truthful histories" $\mathcal{H}(T)$. A history is truthful if m has been played less than $\lfloor pK \rfloor$ times in the current block and M has been played less than $\lfloor (1-p)K \rfloor$ times in the current block. Once either action has been played this many times, histories are deterministic for the rest of the block:

$$H_t \in \begin{cases} \mathcal{H}(T) & \text{if } \#\{\tau | t_0(t) \leq \tau \leq t-1, a_\tau = m\} < \lfloor pK \rfloor \\ & \qquad \text{and } \#\{\tau | t_0(t) \leq \tau \leq t-1, a_\tau = M\} < \lfloor (1-p)K \rfloor \\ \mathcal{H}(D) & \text{otherwise} \end{cases}.$$

For $H_t \in \mathcal{H}(D)$, let $X(H_t)$ be the first deterministic period in the block containing period t, and let N^a be the number of times that action a had been played over the truthful periods in the block:

$$X(H_t) \equiv \max\{\tau \leq t | H_{\tau-1} \in \mathcal{H}(T)\}$$
$$N^a(H_t) \equiv \#\{\tau | t_0(t) \leq \tau < X(H_t), a_\tau = a\}.$$

On $\mathcal{H}(D)$, if $N^M = \lfloor (1-p)K \rfloor$ then let $\underline{a}(H_t) = m$ and $\overline{a}(H_t) = M$; otherwise, if $N^m = \lfloor pK \rfloor$, then let $\underline{a}(H_t) = M$ and $\overline{a}(H_t) = m$. Either $N^M = \lfloor (1-p)K \rfloor$ or $N^m = \lfloor pK \rfloor$ at a deterministic period, because periods only become deterministic once one of these holds. In words, $\overline{a}$ is the action that has been played "enough" over the truthful periods while $\underline{a}$ is the action which "needs more plays".

Now, let all experts of all types share the following conditional strategy. At any history in $\mathcal{H}(T)$ the expert plays truthfully, and at any history in $\mathcal{H}(D)$ the expert plays the underplayed action $\underline{a}(H_t)$: for all e^i, in state is θ^a,

$$\sigma_r^i(H_t, \theta^a) = \begin{cases} a & \text{if } H_t \in \mathcal{H}(T) \\ \underline{a}(H_t) & \text{if } H_t \in \mathcal{H}(D) \end{cases}$$

See Figure 2 for an illustration of the experts' strategy.

Now we construct the customers' strategy ρ. At period 1, c^1 chooses $\rho^1(H_1) = e^1$. At period $t > 1$, if the previous expert chosen was e^i (so that $H_t = H_{t-1} \oplus (e^i, a_{t-1})$), the customer chooses

$$\rho^t(H_t) = \begin{cases} e^i & \text{with Prob } q(H_t) \\ e^{i+1} & \text{with Prob } 1 - q(H_t) \end{cases}$$

where the function $q : \mathcal{H} \setminus \mathcal{H}_1 \to [0, 1]$ determines the probability of continuing with an expert rather than firing her and moving to the next one. $q(H_t)$ satisfies

$$q(H_t) = \begin{cases} q^{StartT}(H_t) & \text{if } t = t_0(t) \;\&\; a_{t-1} = \underline{a}(H_{t-1}) \\ 0 & \text{if } t = t_0(t) \;\&\; a_{t-1} \neq \underline{a}(H_{t-1}) \\ 1 & \text{if } H_t \in \mathcal{H}(T) \;\&\; t > t_0(t) \\ q^{StartD}(H_t) & \text{if } H_t \in \mathcal{H}(D) \;\&\; t = X(H_t) \\ 1 & \text{if } H_t \in \mathcal{H}(D) \;\&\; t > X(H_t) \;\&\; a_{t-1} = \underline{a}(H_t) \\ 0 & \text{if } H_t \in \mathcal{H}(D) \;\&\; t > X(H_t) \;\&\; a_{t-1} \neq \underline{a}(H_t) \end{cases}$$

with q^{StartD} and q^{StartT} defined below. At t_0, the start of a new block – and therefore the end of an old block – there is probability q^{StartT} of keeping the previous expert if she played the suggested action in the previous period, and probability 0 otherwise. At every other truthful period, the customer returns to the previous expert with probability 1 no matter what. At the start of the first deterministic period, customers again move to a new expert with probability $1 - q^{StartD}$. For the remaining deterministic periods in the block, the customer keeps the previous expert with probability 1 if the expert plays the suggested action $\underline{a}$ and fires her otherwise.

It will be useful to define a few other terms on the way to constructing q^{StartD} and q^{StartT}. For a deterministic history H_t, let $Z^a(H_t)$ be the weight that would accumulate towards action a (relative to $t_0(t)$) for an expert chosen at t_0 intending to play actions consistent with H_t over the truthful periods in the block, t_0 through $X(H_t) - 1$. Let $W^a(H_t)$ be the

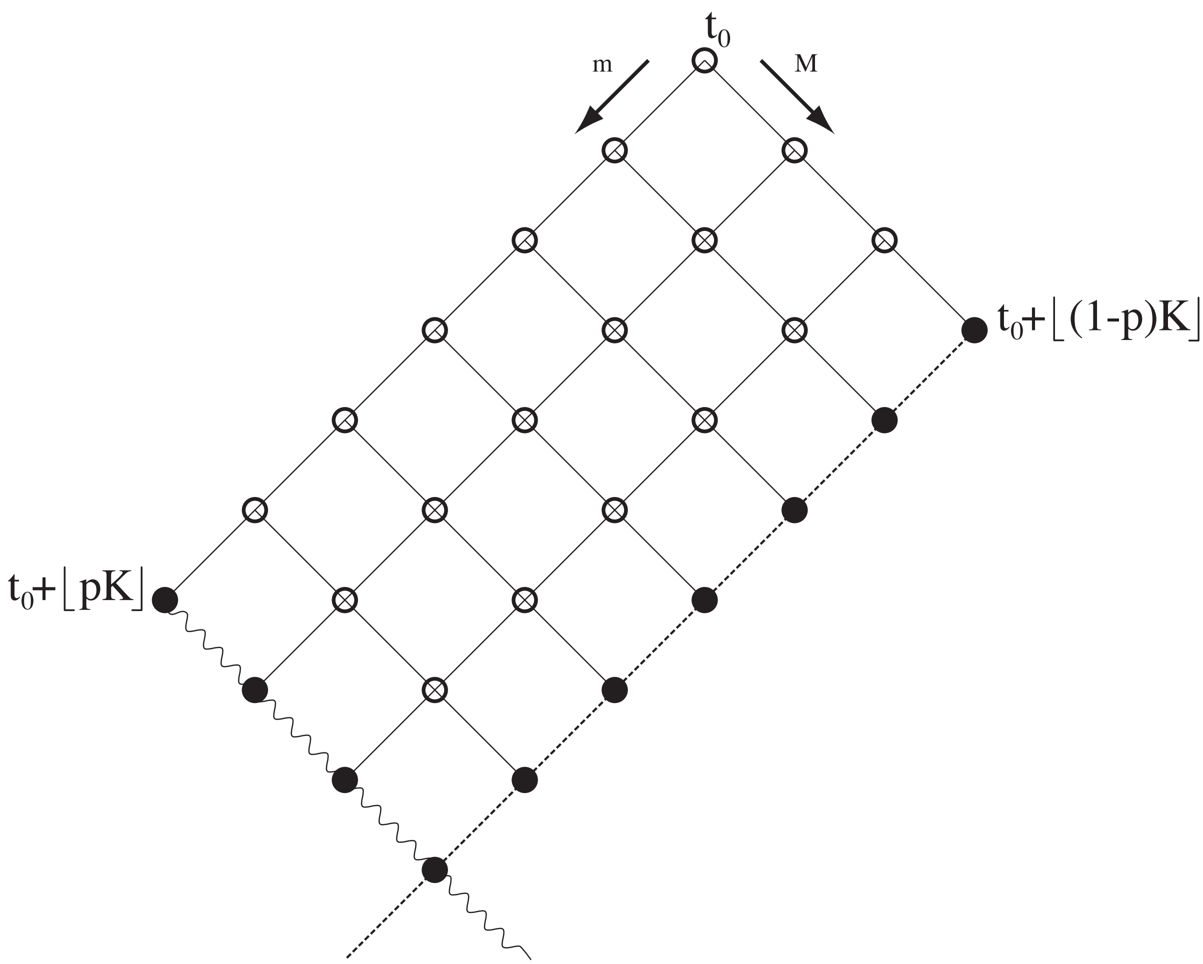

Figure 2: Expert Strategies in Proposition 2, for $K = 10$ and $p = 2/3$.

This picture illustrates the strategy of the experts in Proposition 2 in a single K-period block, for $K = 10$ and $p = 2/3$. At the end of the block, the strategy repeats. The open circles represent truthful periods; the closed circles represent deterministic periods. The style of the lines is varied in order to show which histories lead to which actions at deterministic periods.

On the equilibrium path, the customer chooses the continuation probability as $q = 1$ at every history except for the first deterministic history in a block, at which q^{StartD} is chosen; and the truthful history at t_0, at which q^{StartT} is chosen. These q's depend on the full history of actions at all truthful periods in the most recent block.

In this example, $\lfloor pK \rfloor = 6$ and $\lfloor (1-p)K \rfloor = 3$. The expected number of truthful periods in a block is about 6.60, so the long-term proportion of truthful actions is 66%.

weight that would accumulate over all periods in the block, from t_0 through $t_0 + K - 1$, for an expert chosen at t_0 intending to play actions consistent with H_t over the truthful periods and the equilibrium action $\underline{a}(H_t)$ for the rest of the block. That is, on $H_t \in \mathcal{H}(D)$,

$$Z^a(H_t) \equiv \sum_{\substack{\tau \text{ s.t. } a_\tau = a \\ \& \ t_0(t) \leq \tau < X(H_t)}} \beta^{\tau - t_0(t)}$$

$$W^a(H_t) \equiv \begin{cases} Z^a(H_t) & \text{if } a = \overline{a}(H_t) \\ Z^a(H_t) + q^{StartD}(H_t) \sum_{\tau=X(H_t)}^{t_0(t)+K-1} \beta^{\tau - t_0(t)} & \text{if } a = \underline{a}(H_t) \end{cases}.$$

Adjusting q^{StartD} lets us adjust the weight $W^{\underline{a}}$ that accumulates towards $\underline{a}$ within a single block without affecting the weight $W^{\overline{a}}$ that accumulates towards $\overline{a}$. We want to choose q^{StartD} so that the ratio of weights $\frac{W^m}{W^M}$ is equal to the respective ratio of the probabilities of the actions being appropriate, $\frac{p}{1-p}$, across all equilibrium paths. For $H_t \in \mathcal{H}(D)$, let $q^{StartD}(H_t)$ be defined by

$$q^{StartD}(H_t) \equiv \begin{cases} \frac{\frac{p}{1-p} Z^M(H_t) - Z^m(H_t)}{\sum_{\tau=X(H_t)}^{t_0(t)+K-1} \beta^{\tau - t_0(t)}} & \text{if } \underline{a}(H_t) = m \\ \frac{\frac{1-p}{p} Z^m(H_t) - Z^M(H_t)}{\sum_{\tau=X(H_t)}^{t_0(t)+K-1} \beta^{\tau - t_0(t)}} & \text{if } \underline{a}(H_t) = M \end{cases}.$$

Rearranging, we see that q^{StartD} has been chosen so that

$$\begin{aligned} pW^M(H_t) &= (1-p)W^m(H_t) \quad \text{if } \underline{a} = m \\ (1-p)W^m(H_t) &= pW^M(H_t) \quad \text{if } \underline{a} = M \end{aligned}$$

and so in either case, $\frac{W^m}{W^M} = \frac{p}{1-p}$ as desired.

Now, for $H_t \in \mathcal{H}(D)$, define $Y^a(H_t)$ as

$$Y^a(H_t) \equiv \frac{W^a(H_t)}{1 - \beta^K q^{StartD}(H_t)}$$

and define $\underline{Y}^a$ as

$$\underline{Y}^a \equiv \min_{H_t \in \mathcal{H}(D)} Y^a(H_t).$$

The minimum is well-defined because blocks are identical, and there are only finitely many action paths along the truthful periods of a block. If two deterministic histories share the same action path over the truthful periods in their respective blocks then the histories have identical Y^a values. Y^m is equal to $\frac{p}{1-p} Y^M$ and so $\underline{Y}^m = \frac{p}{1-p} \underline{Y}^M$.

$Y^a(H_t)$ would be the lifetime weight on a, relative to t_0, that an expert chosen at t_0 would receive if she planned to repeat the actions played in the truthful periods of the current block of H_t in the truthful periods of every future block and to play the suggested actions in all deterministic periods, if the continuation probability across blocks were q^{StartD}. But in fact the continuation probability is not q^{StartD} but $q^{StartD} \cdot q^{StartT}$, because the expert may be fired at both the first deterministic period in a block and also at the start of the next block. Just as adjusting q^{StartD} allowed us to manipulate the relative weights on m and M along a path, adjusting q^{StartT} will let us affect the level of the weights along a repeating path while holding the relative weights fixed. We want to set $q^{StartT}(H_t)$ so that the lifetime weight on a is equal to $\underline{Y}^a$ along any repeating path.

For $t_0 \in T_0$ with $t_0 > 1$, let

$$q^{StartT}(H_{t_0}) \equiv \frac{1 - \frac{W^m(H_{t_0-1})}{\underline{Y}^m}}{\beta^K q^{StartD}(H_{t_0-1})} = \frac{1 - \frac{W^M(H_{t_0-1})}{\underline{Y}^M}}{\beta^K q^{StartD}(H_{t_0-1})}.$$

It holds that

$$\underline{Y}^a = \frac{W^a(H_{t_0-1})}{1 - \beta^K q^{StartD}(H_{t_0-1}) q^{StartT}(H_{t_0})}$$

where the right-hand side is the actual lifetime weight on a for an expert picked at t_0 who plans to repeat the path consistent with H_t. This is constant across all H_t. It will turn out that if these levels are constant across repeating paths, they will also be constant across all prospective equilibrium paths.

This completes the descriptions of the strategies. Before we check whether these strategies imply an equilibrium for high discount factors, we need to show that q^{StartD} and q^{StartT} are valid probabilities for β large enough, i.e., that they are numbers in [0,1].

- $\lim_{\beta \to 1} q^{StartD}(H_t) \in (\frac{1}{4} \min\{\frac{p}{1-p}, \frac{1-p}{p}\}, 1)$:

  As $\beta \to 1$,

$$Z^m \to N^m = \begin{cases} (X - t_0) - \lfloor (1-p)K \rfloor & \text{if } \underline{a} = m \\ \lfloor pK \rfloor & \text{if } \underline{a} = M \end{cases}$$
$$Z^M \to N^M = \begin{cases} \lfloor (1-p)K \rfloor & \text{if } \underline{a} = m \\ (X - t_0) - \lfloor pK \rfloor & \text{if } \underline{a} = M \end{cases}.$$

So

$$q^{StartD} = \begin{cases} \dfrac{\frac{p}{1-p}Z^M - Z^m}{\sum_{\tau=X(H_t)}^{t_0(t)+K-1} \beta^{\tau-t_0(t)}} \xrightarrow{\beta\to 1} \dfrac{\frac{p}{1-p}\lfloor(1-p)K\rfloor - (X-t_0) + \lfloor(1-p)K\rfloor}{K-(X-t_0)} = \dfrac{\frac{1}{1-p}\lfloor(1-p)K\rfloor - (X-t_0)}{K-(X-t_0)} & \text{if } \underline{a} = m \\ \dfrac{\frac{1-p}{p}Z^m - Z^M}{\sum_{\tau=X(H_t)}^{t_0(t)+K-1} \beta^{\tau-t_0(t)}} \xrightarrow{\beta\to 1} \dfrac{\frac{1-p}{p}\lfloor pK\rfloor - (X-t_0) + \lfloor pK\rfloor}{K-(X-t_0)} = \dfrac{\frac{1}{p}\lfloor pK\rfloor - (X-t_0)}{K-(X-t_0)} & \text{if } \underline{a} = M \end{cases}.$$

Because pK is not an integer, in either case the numerator is strictly smaller than the (positive) denominator; the limit of q^{StartD} is strictly less than 1.

Now we wish to show that the minimum value of $\lim q^{StartD}$ over all deterministic histories is greater than $\frac{1}{4}\min\{\frac{p}{1-p}, \frac{1-p}{p}\}$.

The maximum value that $(X - t_0)$ can take is $\lfloor pK\rfloor + \lfloor(1-p)K\rfloor - 1 = K - 2$. And for any fixed $\underline{a}$, the above expression for $\lim q^{StartD}$ is decreasing in $(X - t_0)$ as long as $(X - t_0) < K$. This implies that for any $H_t \in \mathcal{H}(D)$,

$$\lim q^{StartD}(H_t) \geq \min \left\{ \frac{\frac{1}{1-p}\lfloor(1-p)K\rfloor - K + 2}{2}, \frac{\frac{1}{p}\lfloor pK\rfloor - K + 2}{2} \right\}.$$

Noting that $\lfloor(1-p)K\rfloor = K - \lfloor pK\rfloor - 1$, the first fraction can be reduced to

$$\frac{\lfloor(1-p)K\rfloor - (1-p)K + 2(1-p)}{2(1-p)} = \frac{pK - \lfloor pK\rfloor - 2p + 1}{2(1-p)} > \frac{\zeta}{2(1-p)}$$

and the second can be reduced to

$$\frac{\lfloor pK\rfloor - pK + 2p}{2p} > \frac{\zeta}{2p}$$

where the inequalities come from condition (iii) of Lemma 3, with $\zeta = \min\{p/2, (1-p)/2\}$.

Therefore $\lim q^{StartD}(H_t)$ is greater than $\min\{\frac{\zeta}{2(1-p)}, \frac{\zeta}{2p}\} = \frac{1}{4}\min\{\frac{p}{1-p}, \frac{1-p}{p}\}$.

- $\lim_{\beta\to 1} q^{StartT}(H_t) \in (0, 1]$:

  For $t = t_0 - 1$,

$$q^{StartT}(H_{t+1}) = \frac{1 - \frac{W^m(H_t)}{\underline{Y}^m}}{\beta^K q^{StartD}(H_t)} \leq \frac{1 - \frac{W^m(H_t)}{Y^m}}{\beta^K q^{StartD}(H_t)} = \frac{1 - \frac{W^m(H_t)}{\frac{W^m(H_t)}{1-\beta^K q^{StartD}(H_t)}}}{\beta^K q^{StartD}(H_t)} = 1$$

  and so $q^{StartT} \leq 1$. Now we will show that the limit of $q^{StartT}(H_{t+1})$ as β goes to 1 is strictly positive.

$$\lim q^{StartT}(H_{t+1}) = \frac{1 - \frac{\lim W^m(H_t)}{\lim \underline{Y}^m}}{\lim q^{StartD}(H_t)} = \frac{1 - \frac{\lim W^m(H_t)}{\frac{\hat{W}^m}{1-\hat{q}^{StartD}}}}{\lim q^{StartD}(H_t)}$$

where $\hat{q}^{StartD} = \lim q^{StartD}(\hat{H})$ and $\hat{W}^m = \lim W^m(\hat{H})$ for some $\hat{H} \in \mathcal{H}(D)$ with $\lim Y^m(\hat{H}) = \lim \underline{Y}^m$. This is positive if and only if

$$\frac{\hat{W}^m}{\lim W^m(H_t)} > 1 - \hat{q}^{StartD}.$$

We know that $\hat{q}^{StartD} > \frac{1}{4}\min\{\frac{p}{1-p}, \frac{1-p}{p}\}$ so it suffices to show that

$$\frac{\hat{W}^m}{\lim W^m(H_t)} > 1 - \frac{1}{4}\min\{\frac{p}{1-p}, \frac{1-p}{p}\}. \tag{1}$$

Notice that $\lim W^m(H_t)$ can be expressed as

$$\lim W^m(H_t) = \begin{cases} N^m(H_t) = \lfloor pK \rfloor & \text{if } \underline{a} = M \\ \frac{p}{1-p} N^M(H_t) = \frac{p}{1-p}\lfloor (1-p)K \rfloor & \text{if } \underline{a} = m \end{cases}$$

and the same holds for $\hat{W}^m$. Therefore (1) follows from condition (iv) of Lemma 3.

So for β sufficiently large, we have defined a valid strategy profile. It remains to be shown that this strategy profile is an equilibrium when β is close to 1, and that the proportion of truthfulness in this equilibrium goes to 1 as K increases.

Because all experts act identically at every history, on or off the equilibrium path, any customer strategy will be a best response. So to show that the strategy is an equilibrium, it will suffice to show that the experts play best responses at every history when β is large enough.

By the following lemma, to show that the strategy is a best response for experts at truthful periods, we only need to look at how deviations at later periods would affect the weights relative to t_0. If an expert chosen at t_0 would get the same set of t_0-weights from planning to deviate to m as she would deviating to M at any later truthful period in the block, then there will be no profitable deviation once any such period is reached.

Lemma 4. Let $\sigma_{H_\tau,a} : \mathcal{H} \times \Theta \to \Delta(A)$ for $a \in A, H_\tau \in \mathcal{H}$ be identical to the conditional strategy for σ^i_τ, defined above, except at the history H_τ. At H_τ, $\sigma_{H_\tau,a}$ plays a for either θ. Take some $t_0 \in T_0$ to be the start of a block, and suppose that for all $H_\tau \in \mathcal{H}(T)$ satisfying $t_0(\tau) = t_0$ (that is, for all truthful histories in that block) it holds that $W^i_a(H_{t_0}, \sigma_{H_\tau,m}) = W^i_a(H_{t_0}, \sigma_{H_\tau,M})$ for $a = m, M$. Then e^i has no profitable deviation if selected at any $H_\tau \in \mathcal{H}(T)$ satisfying $t_0(\tau) = t_0$.

Proof. See Appendix. □

We will show that the conditions of Lemma 4 hold, implying that our strategy is in fact a best response at truthful periods.

If expert e^i is picked at time $t_0 \in T_0$, she receives weights $W^i_m(H_{t_0}, \sigma^i), W^i_M(H_{t_0}, \sigma^i)$ from following the proposed strategy. Consider a path of actions from t_0 to $t_0 + K - 1$, going from the first period in the block to the last period before the block repeats. Relative to t_0, a weight of $W^a(H_{t_0+K-1})$ accumulates towards $W^i_a(H_{t_0}, \sigma^i)$. Because strategies repeat anew every K periods, $W^i_a(H_{t_0}, \sigma^i) = W^i_a(H_{t_0+K}, \sigma^i)$, and so $W^i_a(H_{t_0}, \sigma^i)$ satisfies the recursive formula

$$W^i_a(H_{t_0}, \sigma^i) = \mathbb{E}\Big[W^a(H_{t_0+K-1}) + \beta^K q^{StartD}(H_{t_0+K-1}) q^{StartT}(H_{t_0+K}) W^i_a(H_{t_0}, \sigma^i) \Big| H_{t_0}\Big]$$

where the expectation is taken over H_{t_0+K}, given H_{t_0}. On each path, $q^{StartT} = \frac{1 - \frac{W^a}{\underline{Y}^a}}{q^{StartD}\beta^K}$. Plugging this in to the recursive formula gives

$$\begin{aligned}
W^i_a(H_{t_0}, \sigma^i) &= \mathbb{E}\Big[W^a(H_{t_0+K-1}) + \Big(1 - \tfrac{W^a(H_{t_0+K-1})}{\underline{Y}^a}\Big) W^i_a(H_{t_0}, \sigma^i) \Big| H_{t_0}, \sigma^i\Big] \\
\Longrightarrow \qquad 0 &= \Big(1 - \tfrac{W^i_a(H_{t_0}, \sigma^i)}{\underline{Y}^a}\Big) \mathbb{E}\Big[W^a(H_{t_0+K-1}) | H_{t_0}, \sigma^i\Big] \\
\Longrightarrow \quad W^i_a(H_{t_0}, \sigma^i) &= \underline{Y}^a.
\end{aligned}$$

Now consider a deviation of the form discussed in Lemma 4 to the strategy $\sigma^i_{H_\tau,a}$. Following the same substitutions, with $W^i_a(H_{t_0}, \sigma^i) = \underline{Y}^a$, this gives weights

$$\begin{aligned}
&W^i_a(H_{t_0}, \sigma^i_{H_\tau,a}) \\
&\quad = \mathbb{E}\Big[W^a(H_{t_0+K-1}) + \beta^K q^{StartD}(H_{t_0+K-1}) q^{StartT}(H_{t_0+K}) W^i_a(H_{t_0}, \sigma^i) \Big| H_{t_0}, \sigma^i_{H_\tau,a}\Big] \\
&\quad = W^i_a(H_{t_0}, \sigma^i) + \left(1 - \frac{W^i_a(H_{t_0}, \sigma^i)}{\underline{Y}^a}\right) \mathbb{E}\Big[W^a(H_{t_0+K-1}) | H_{t_0}, \sigma^i_{H_\tau,a}\Big] \\
&\quad = W^i_a(H_{t_0}, \sigma^i)
\end{aligned}$$

(In equilibrium, if expert e^i is selected at t_0 then the same expert will be selected at every truthful period in a block, and W^a, q^{StartD}, and q^{StartT} are determined only by the play at truthful periods. So when the expectation is taken, only e^i's strategy need be considered.). So all such deviations yield the same point in (W_m, W_M)-space, and Lemma 4 applies; there are no profitable deviations by experts at any $\mathcal{H}(T)$ periods.

To finish the proof that the strategies are an equilibrium for β and K large enough, we now only need to show that the expert prefers not to deviate at $\mathcal{H}(D)$ periods. Starting at any $\mathcal{H}(D)$ period H_t, the expert is at most K periods away from reaching the next block at $t_0(t+K)$ with probability q^{StartT}. So from following the equilibrium path, $W_a^i(H_t, \sigma^i) \geq \beta^K q^{StartT} W_a^i(H_{t_0(t+K)}, \sigma^i) = \beta^K q^{StartT} \underline{Y}^a$. Fixing K and taking $\beta \to 1$, β^K goes to 1 and q^{StartT} approaches some value at least equal to $\min\{\frac{\zeta}{2(1-p)}, \frac{\zeta}{2p}\}$. The limit of $\underline{Y}^m$ is bounded below by an expression which goes to infinity:

$$\lim_{\beta\to 1} \underline{Y}^m \geq \min_{H_\tau \in \mathcal{H}(D)} \lim_{\beta \to 1} W^m(H_\tau) \geq \min\left\{pK - 1, \frac{p}{1-p}((1-p)K - 1)\right\} = pK - \max\left\{1, \frac{p}{1-p}\right\}$$

and a similar argument shows that $\underline{Y}^M$ also diverges. Therefore, for β and K large enough, $W_a^i(H_{t_0}, \sigma^i)$ becomes arbitrarily large and in particular is greater than 1. Deviating from $\underline{a}$ to $\overline{a}$ at $H_t \in \mathcal{H}(D)$ gives a weight of 1 on $\overline{a}$ and 0 on $\underline{a}$, and so for K and β large enough this is strictly dominated by not deviating because $r^i \geq 0$.

So the strategies we have constructed do form an equilibrium. Now we show that as K increases, these strategies give an arbitrarily high long-term proportion of truthful play.

Consider the probability of having less than or equal to n truthful periods in a block of length K. Writing p_a as the probability of state θ^a in a period, so $p_m = p$ and $p_M = 1 - p$, and letting x_a^n denote a binomial random variable of n draws from probability p_a,

$$\begin{aligned} Prob\big[\text{At most } n \text{ truthful periods in a block}\big] &= Prob\big[\text{At least } \lfloor p_a K \rfloor\ \theta^a\text{'s after } n \text{ periods, for some } a\big] \\ &\leq \sum_a Prob\big[\text{At least } \lfloor p_a K \rfloor\ \theta^a\text{'s after } n \text{ periods}\big] \\ &= \sum_a Prob\big[x_a^n \geq \lfloor p_a K \rfloor\big] \\ &\leq \sum_a Prob\big[\frac{x_a^n}{n} \geq p_a + p_a \frac{K-n-1}{n}\big]. \qquad (2) \end{aligned}$$

The random variable x_a^n/n has mean p_a and standard deviation $\sqrt{\frac{p_a(1-p_a)}{n}}$. So if we take $n \simeq K - K^s$ for some $s \in (1/2, 1)$ then $p_a \frac{K-n-1}{n}$ divided by the standard deviation goes to infinity as K goes to infinity:

$$\frac{p_a \frac{K-n-1}{n}}{\sqrt{\frac{p_a(1-p_a)}{n}}} \simeq \frac{p_a \frac{K^s-1}{K-K^s}}{\sqrt{\frac{p_a(1-p_a)}{K-K^s}}} = \sqrt{\frac{p_a}{1-p_a}} \frac{K^s-1}{\sqrt{K-K^s}} \geq \sqrt{\frac{p_a}{1-p_a}} \frac{K^s-1}{\sqrt{K}} \underset{K\to\infty}{\longrightarrow} \infty.$$

Therefore each of the probability terms in (2) goes to 0 as K increases (for instance, by Chebyshev's theorem), and the probability of more than $n \simeq K - K^s$ out of K truthful periods goes to 1. Moreover, since $\frac{n}{K} \simeq \frac{K-K^s}{K}$ goes to 1 for large enough K, the expected proportion of truthful periods in a block must approach 1 for large enough K.[4]

Because blocks are independent, the Law of Large Numbers tells us that the long-term proportion of truthful periods approaches the expected proportion in a given block, and we can get this arbitrarily close to 1. □

5 Extensions

We consider an extreme environment in which customers are short term players who never receive signals about the true state in past periods, and in which expert preferences are completely independent of the state of the world. Any combination of these assumptions can be relaxed without fundamentally altering our conclusions.

The stream of customers can be thought of as a single long-term player, or some combination of short- and long-term players, without affecting any of the equilibria. The customers' actions have no effect on the current or future play of experts, and so any strategy is both myopically and dynamically optimal. We focussed on short-lived customers to highlight the fact that long-term relationships between individuals are unnecessary so long as histories are observable.

If experts receive a small amount of disutility from mismatching the action and the state – due to guilt from lying, a fear of God or audits, or because their underlying cost structure depends on the state – then all equilibrium strategy profiles in the paper remain equilibria. Indeed, a slight preference for truth makes truthfulness a strict rather than a weak best response at the appropriate periods.

Moreover, because all of the equilibria we construct are fully pooling, signals about the true state in past periods reveal no new information about an expert's type or about how

[4]This implies that the number of deterministic periods in a block is of order at most $\sqrt{K}$; taking $s < 1/2$, the difference divided by the standard deviation goes to 0, which can be used to show that the order is at least $\sqrt{K}$.

an expert will play in the future. Signals may alter the set of equilibria, but do not disrupt the ones we lay out. (It is easy to imagine that long-lived customers in particular might observe signals of an expert's past truthfulness – for instance, a customer brought his car into a mechanic to have it repaired, and the car still had problems when he got it back).

The assumption that the outside option of an expert is 0 may not make sense if we interpret experts as having no business outside of this market. How could we possibly support a large (infinite) number of experts, when almost all of them get no business? But if we think that the experts have nonbinding capacity constraints and otherwise linear costs, then this constructed market can be thought of as being on top of whatever other business they get – possibly from identical markets running in parallel.

In this model, the infinity of available experts stands in for the ability of a customer to go to a new expert in each period. We can implement all of the equilibria considered so long as the previously chosen expert and a single new one are always available. In the Web Appendix we show how such an equilibrium can work in an environment akin to Section 3 in which customers must always return to a fixed, finite set of experts. We also extend Section 3 to a setting with a larger action space in which there may be observable heterogeneity.

Finally, in the Web Appendix we discuss the possibility of truthful play in a one-shot setting where prices may not be exogenously fixed. When experts are homogeneous or when their costs are commonly known, prices for each treatment can be found at which experts earn the same profit for each action. At these prices, experts will be truthful even when no future business depends on the action chosen today. But it is impossible to find such prices when an expert's costs are drawn from a nondegenerate distribution and are privately observed. There is no way to condition prices on reported costs appropriately without giving experts an incentive to lie about these costs.

6 Conclusion

We looked at a model where short-lived customers successively choose long-lived experts who decide on and then perform actions for the customers. Customer utilities depend on the state of the world along with the chosen action, but customers cannot observe the state. Customers only observe the history of past actions taken by experts. The experts see the state, but it doesn't factor into their utility function. We considered how experts could be induced to take the state-dependent action preferred by the customers.

In the game where experts do not have private information about their own preferences, a truthful equilibrium can be implemented in the following manner. If the previous period's expert just performed a minor treatment, then the next customer returns to that expert; if the expert just performed a major one then the customer moves to a new expert with some probability. The expert is indifferent between the actions because she gets more money today but less future business from a major treatment, and the customer is indifferent across experts because all would be truthful.

When experts do have private information about their own preferences, fully truthful equilibria are no longer possible. But customers can play a strategy in which all types of experts will be indifferent over actions, and will therefore play truthfully, in certain periods. In other periods, the experts are told to ignore the state and perform some predetermined action. At the truthful periods, experts are indifferent because either action will lead to the expert's performing the same number of expected, discounted lifetime minor treatments and the same number of expected, discounted lifetime major treatments. As the discount factor approaches 1, we can achieve truthful actions in nearly all periods.

Appendix: Proofs of Lemmas

Proof of Lemma 1. The strategy σ^i is sequentially rational and gives an optimal payoff starting from every history, so it is sufficient to show that $\hat{\sigma}^i$ gives an optimal payoff starting from any history in $\hat{\mathcal{H}}$.

We divide this into three cases.

- **Case 1:** There is a single history $\hat{H}_t \in \hat{\mathcal{H}}$. The strategy of playing m at $\hat{H}_t$ and continuing with σ^i in the future gives the same payoff as the strategy of playing M at $\hat{H}_t$ followed by σ^i; otherwise, it would not be optimal for e^i to be truthful. So for any realization of θ_t, any mixture of these two strategies also gives this same optimal payoff.

- **Case 2:** $\hat{\mathcal{H}}$ is finite. We can apply the argument of Case 1 inductively, changing the strategies at each element of $\hat{\mathcal{H}}$ in any order. After each of these changes, the strategy remains sequentially rational and payoffs remain the same.

- **Case 3:** $\hat{\mathcal{H}}$ is countably infinite. Suppose that $\hat{\sigma}^i$ gives $\delta > 0$ less utility than σ^i to e^i if she is chosen at some history H_t. For any positive N, we can change the strategy from σ^i to $\hat{\sigma}^i$ at the finitely many histories $\hat{H}_\tau \in \hat{\mathcal{H}}$ which satisfy $t \leq \tau \leq t+N$, and utilities at all periods will remain constant as in Case 2. Call this intermediate strategy σ^i_N.

  The highest stage payoff that the player can receive is 1 and the lowest is 0, and so starting at H_t the utilities from strategies $\hat{\sigma}^i$ and σ^i_N can differ by at most $\sum_{\tau=N+1}^{\infty} \beta^\tau =$

$\frac{\beta^{N+1}}{1-\beta}$. For N large enough, this difference must be less than any fixed $\delta > 0$. Contradiction.[5] □

Proof of Lemma 2. The "If" part is immediate from the fact that, given on an expert's type, the utility of a conditional strategy is determined entirely by the weights it induces on M and m.

To show the "Only If" part, let e^i be indifferent between σ' and σ'' at history H_t for both r' and r''. Then

$$\begin{aligned} &\begin{cases} r'W_m^i(H_t,\sigma') + W_M^i(H_t,\sigma') = r'W_m^i(H_t,\sigma'') + W_M^i(H_t,\sigma'') \\ r''W_m^i(H_t,\sigma') + W_M^i(H_t,\sigma') = r''W_m^i(H_t,\sigma'') + W_M^i(H_t,\sigma'') \end{cases} \\ \implies &\begin{cases} r'\big(W_m^i(H_t,\sigma') - W_m^i(H_t,\sigma'')\big) = W_M^i(H_t,\sigma'') - W_M^i(H_t,\sigma') \\ r''\big(W_m^i(H_t,\sigma') - W_m^i(H_t,\sigma'')\big) = W_M^i(H_t,\sigma'') - W_M^i(H_t,\sigma') \end{cases} \\ \implies &(r'-r'')\big(W_m^i(H_t,\sigma') - W_m^i(H_t,\sigma'')\big) = 0. \end{aligned}$$

Because $r' \neq r''$, it must hold that $W_m^i(H_t,\sigma') = W_m^i(H_t,\sigma'')$. Plugging this back into the original indifference $r'W_m^i(H_t,\sigma') + W_M^i(H_t,\sigma') = r'W_m^i(H_t,\sigma'') + W_M^i(H_t,\sigma'')$ implies that $W_M^i(H_t,\sigma') = W_M^i(H_t,\sigma'')$ as well. □

Proof of Lemma 3. Conditions (i) and (iv) hold for any K large enough. Condition (ii) holds if $pK - \lfloor pK \rfloor \neq 0$.

Condition (iii) is equivalent to $pK - \lfloor pK \rfloor \in (\frac{5p-2}{2}, \frac{3p}{2})$ for $p \leq \frac{1}{2}$, or $pK - \lfloor pK \rfloor \in (\frac{3p-1}{2}, \frac{5p-1}{2})$ for $p \geq \frac{1}{2}$. In either case, both conditions (ii) and (iii) will be satisfied if $pK - \lfloor pK \rfloor$ is in some small neighborhood $\mathcal{N}_p \subseteq (0,1)$ about p. (In fact, even for K small, condition (iv) is satisfied if $pK - \lfloor pK \rfloor$ is close to p).

For $p \in (0,1)$ irrational, $\{pn - \lfloor pn \rfloor | n \in \mathbb{N}\}$ is dense on $(0,1)$. For p rational with reduced denominator d, any K of the form $K = nd + 1$ will have $pK - \lfloor pK \rfloor = p \in \mathcal{N}_p$. In each case an arbitrarily large K can be found with $pK - \lfloor pK \rfloor$ in $\mathcal{N}_p$. □

Proof of Lemma 4. Let $\mathcal{G}_\tau = \{H_\tau \in \mathcal{H}(T)\}$ be the set of possible truthful histories at time τ. Say that H'_τ and H''_τ in $\mathcal{G}_\tau$ are *equivalent* if the actions from periods $t_0(\tau)$ through $\tau - 1$ are the same in both histories. For $\tau = t_0(\tau)$, all H_τ are equivalent. The experts' and customers' strategies are such that picking an expert e^i at two equivalent histories yields identical play by all agents going forward.

Take $t_0 \in T_0$, and fix some τ such that $t_0(\tau) = t_0$ and $\mathcal{G}_\tau$ is nonempty. Suppose that e^i is selected at H_t. Define $G_\tau^i(H_\tau | H_{t_0})$ to be the single history following H_{t_0} that is equivalent to H_τ, in which e^i is selected at period t_0 and is selected with positive probability at period τ according to the customers' equilibrium strategy. It is the element of the equivalence class following H_{t_0} in which e^i is chosen in every period from t through $\tau - 1$ and the actions

[5] When this lemma is applied in Section 4, the highest stage payoff is $\max\{r^i, 1\}$ rather than 1, and so the maximum payoff difference is $\frac{\beta^{N+1}}{1-\beta} \cdot \max\{r^i, 1\}$ rather than $\frac{\beta^{N+1}}{1-\beta}$. The argument is otherwise unchanged.

corresponding to H_τ are played. For $\tau = t_0$, define $G^i_\tau(H_\tau|H_{t_0})$ to be H_τ even if e^i is not selected by the customers with positive probability at H_{t_0}.

Conditional on customer c^{t_0} selecting expert e^i at history H_{t_0}, given some arbitrary conditional strategy σ' of e^i, denote the probability of a history H_τ occurring and e^i being selected at H_τ by $\pi^i(H_\tau, H_{t_0}|\sigma')$. For any history $H_\tau \neq G^i_\tau(H_\tau|H_{t_0})$, the customers' strategy is such that either H_τ occurs with probability 0 or e^i is selected with probability 0 at H_τ; in either case, $\pi^i(H_\tau, H_{t_0}|\sigma') = 0$ for any σ'. Under the equilibrium conditional strategy σ^i_r, the probability $\pi^i(G^i_\tau(H_\tau|H_{t_0}), H_{t_0}|\sigma^i_r)$ is in fact positive. An expert chosen at t_0 is never fired before the start of deterministic periods in a block, so $\pi^i(G^i_\tau(H_\tau|H_{t_0}), H_{t_0}|\sigma^i_r)$ is just the probability that the states θ_{t_0} through $\theta_{\tau-1}$ are such that truthful play gives the correct sequence of actions.

Now, consider some conditional strategy σ' which differs from the equilibrium strategy σ^i_r only at H_τ. We can see that $W^i_a(H_{t_0}, \sigma')$ is equal to some constant (the weight added along all histories which do not follow H_τ) plus $\beta^{\tau-t_0}\pi(G^i_\tau(H_\tau|H_{t_0}), \sigma^i_r)$ times a convex combination of $W^i_a(H_\tau, \sigma_{H_\tau,m})$ and $W^i_a(H_\tau, \sigma_{H_\tau,M})$. The convex combination places a weight on $W^i_a(H_\tau, \sigma_{H_\tau,a'})$ equal to the probability (unconditional on θ_τ) of e^i playing a' at H_τ if chosen.

In particular, if $\sigma' = \sigma_{H_\tau,a'}$ then the convex combination places a weight of 1 on $W^i_{H_\tau}(a, \sigma_{H_\tau,a'})$ and 0 on $W^i_a(H_\tau, \sigma_{H_\tau,a''})$, for $a'' \neq a'$.

Suppose that the condition of the lemma holds: $W^i_a(H_{t_0}, \sigma_{H_\tau,m}) = W^i_a(H_{t_0}, \sigma_{H_\tau,M})$. Then it must be the case that either $\pi(H_\tau, H_{t_0}|\sigma^i_r) = 0$, or $W^i_a(H_\tau, \sigma_{H_\tau,m}) = W^i_a(H_\tau, \sigma_{H_\tau,M})$. These weights are equal for equivalent histories, and each history is equivalent to one for which π is positive. So in fact $W^i_a(H_\tau, \sigma_{H_\tau,m}) = W^i_a(H_\tau, \sigma_{H_\tau,M})$ for all $H_\tau \in \mathcal{H}(T)$.

Therefore at the history H_τ, if e^i is selected, she has no profitable deviations. She can either deviate to m, M, or some mixture of the two; and any such deviation yields the same weights on each action, that is, the same number of expected discounted lifetime plays. Any such mixture of actions is optimal at a truthful period H_τ. $\square$

References

Ingela Alger and Francois Salanie. A theory of fraud and overtreatment in experts markets. *Journal of Economics & Management Strategy*, 15(4):853–881, Winter 2006.

Heski Bar-Isaac and Steve Tadelis. Seller reputation. *Foundations and Trends in Microeconomics*, 4(4):273–351, 2008.

John A. Byrne. Goodbye to an ethicist. *Business Week*, page 38, February 10, 2003.

Michael R. Darby and Edi Karni. Free competition and the optimal amount of fraud. *Journal of Law and Economics*, 16(1), 1973.

Uwe Dulleck and Rudolf Kerschbamer. On doctors, mechanics and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 44:5–42, 2006.

Uwe Dulleck and Rudolf Kerschbamer. Experts vs. discounters: Consumer free-riding and experts withholding advice in markets for credence goods. *International Journal of Industrial Organization*, 2008.

Jeffey C. Ely and Juuso Valimaki. Bad reputation. *Quarterly Journal of Economics*, 118(3): 785–814, 2003.

Jeffrey C. Ely, Drew Fudenberg, and David K. Levine. When is reputation bad. *Games and Economic Behavior*, 63:498–526, 2008.

Winand Emons. Credence goods and fraudulent experts. *RAND Journal of Economics*, 28 (1):107–119, Spring 1997.

Drew Fudenberg and David K. Levine. Efficiency and observability in games with long-run and short-run players. *Journal of Economic Theory*, 62:103–135, 1994.

Matthew O. Jackson and Hugo Sonnenschein. Overcoming incentive constraints. *Econometrica*, 75(1):241–257, January 2007.

In-Uck Park. Cheap talk referrals of differentiated experts in repeated relationships. *RAND Journal of Economics*, 36(2), Summer 2005.

Martin Pesendorfer. A study of collusion in first price auctions. *Review of Economic Studies*, 67:381–411, 2000.

Carolyn Pitchik and Andrew Schotter. Honesty in a model of strategic information transmission. *American Economic Review*, 77(5):1032–1036, December 1987.

Henry Schneider. Agency problems and reputation in expert services: evidence from auto repair. Johnson School Research Paper Series #15-07, October 2007.

Robert M. Townsend. Optimal multiperiod contracts and the gain from enduring relationships under private information. *The Journal of Political Economy*, 90(6):1166–1186, December 1982.

Asher Wolinsky. Competition in a market for informed experts' services. *RAND Journal of Economics*, 24:380–398, 1993.