Games and Economic Behavior 142 (2023) 150–178

Contents lists available at ScienceDirect

Games and Economic Behavior

journal homepage: www.elsevier.com/locate/geb

Bad reputation with simple rating systems

Caio Lorecchio [a,*], Daniel Monte [b,c,d]

[a] Universitat de Barcelona, School of Economics, John Maynard Keynes 1-11, 08034, Barcelona, Spain
[b] Università di Torino, ESOMAS Department, Corso Unione Sovietica, 218 bis - 10134, Turin, Italy
[c] Collegio Carlo Alberto, Vincenzo Arbarello 8, 10122 Turin, Italy
[d] Fundação Getúlio Vargas, São Paulo School of Economics, Itapeva 474, 01302000, São Paulo, Brazil

ARTICLE INFO

Article history:
Received 15 February 2022
Available online 11 August 2023

JEL classification:
D82
D83
L15
M3

Keywords:
Bad reputation
Rating systems
Bounded rationality

ABSTRACT

We consider information censoring through finite memory as a device against bad reputational concerns. Our class of constrained information policies resembles common practices in online reputation systems, on which customers increasingly rely whenever hiring experts. In a world of repeated interactions between a long-lived expert and short-lived customers, Ely and Välimäki (2003) show that unlimited record-keeping may induce the expert to overchoose a certain action, seeking reputational gains. Consequently, welfare may reduce and markets may break down. We show that simple rating systems in such world help overcome market failures and improve upon both the full-memory and the no-memory cases.

1. Introduction

People rely on online reputation systems for a variety of daily activities. Rating systems help consumers choose their goods, but also help fulfill clients' requests for expert services. For instance, before visiting a doctor's office, patients can check physicians' ratings and reviews at RateMDs. Students can select courses according to teachers' evaluations from other students at Rate My Professors or Rate My Teachers. Americans can hire, rate, and review home service providers on Angi, and motorists can find the best-rated mechanics all around the world on Yelp.

A great deal is known about designing rating systems that rate products, but much less is known about systems that rate economic agents. The main difference between the two classes of rating systems is that, in the latter, the expert being rated can strategically influence his rating, so the design must account for such strategic effects as well.

For concreteness, consider the case of an expert who faces an online feedback record. Some clients might imperfectly ascertain their true needs and evaluate the expert's service negatively, even if the diagnosis and the proposed treatment are correct. This is often true when physicians propose a high-cost intervention for a patient who feels that a lower-cost treatment would lead to the same result, or when teachers are badly evaluated for being rigorous.[1] This imperfect assessment by clients can generate a perverse effect because the expert might have incentives to provide wrong solutions

* Corresponding author.
E-mail addresses: paeslemelorecchio@ub.edu (C. Lorecchio), daniel.monte@unito.it (D. Monte).

[1] Some companies even specialized in helping doctors conceal patients' reviews and devise "anti-ratings" contracts, in which patients are asked to sign away their right to provide information of a medical service online (Goldman, 2010). In France, a legal decision banned Rate My Professors and related sites from naming national teachers (https://bit.ly/3nBoMYO).

https://doi.org/10.1016/j.geb.2023.07.019
0899-8256/

in exchange for a better review. Another example, a mechanic could propose a low-cost tune-up for a driver to avoid being reviewed as an expensive professional, when the most adequate solution would be a costly engine replacement. Additionally, a customer might look at past feedback as deceptive, because the platform could censor information or benefit premium service providers.[2] For those reasons, the benefits of such mediated interactions could be limited.

In this paper, we study the design of rating systems in a reputation game. We consider a stylized interaction between a long-run expert and a sequence of short-run customers. In each period, a customer has the option of hiring the expert's services. If hired, the expert observes a problem and proposes a treatment. A severe problem requires a high-cost treatment and a mild one is solvable through a low-cost treatment. The customer, however, cannot determine how severe the problem is. Moreover, each customer has a common prior belief that the expert is a "bad" commitment type that always provides the expensive treatment.

Our benchmark setting is the bad reputation model by Ely and Välimäki (2003), but with an intermediary committed to a special class of information policy. We chose this model as a benchmark to isolate how the design of a rating system interacts with its strategic implications in a reputation game. In this game, reputational concerns have a striking effect on the outcome.

In the model of Ely and Välimäki (2003), the reputational concern of the expert becomes so strong that the market collapses and customers only rarely hire him. Due to this bad reputation effect, information censoring might improve efficiency. In fact, the outcomes of the game with full censoring - essentially an infinite repetition of the same static game - Pareto dominates the outcomes of the game with no censoring at all (full information), for all players.

The class of information policies we consider consists of (i) a finite message space, (ii) an initial distribution over messages, and (iii) a Markov transition rule, mapping the current message and any possible observable outcome (high-cost treatment, low-cost treatment or no hiring) to a probability distribution over the message space. Technically, information policies are finite automata. Empirically, one can think of them as rating systems, commonly observed in many online reputation systems.

There are several advantages of using finite automata in the construction of rating systems. Online platforms deal with massive amounts of data, so designing a simple rating system overcomes this complexity issue. Additionally, they can be implemented efficiently, allowing for fast and efficient rating systems. Finally, because of the Markov property, finite automata are more resistant to noise and errors, making them suitable for use in rating systems where reliability is important. Therefore, from a computational point of view, finite automata can lead to simpler, more efficient, and more robust systems that are easier to build and maintain.

More importantly, rating systems modeled as finite automata allow for greater accountability since past information is irrelevant for predicting future ratings. Furthermore, such rating systems respect privacy since they do not collect user data apart from the rating itself, which is publicly seen by customers. Thus, these rating systems comply with the current demands of both regulators and end-users for greater accountability and privacy.[3] Our paper helps understand how rating systems subject to these constraints are impacted.

Since we are restricting the class of policies a platform can design, there will be a constraint on what the intermediary can achieve with a rating system. We characterize the upper bounds on the expert's and customers' equilibrium payoffs and construct rating systems that reach such upper bounds.

We show that, to maximize customers' value from the interaction (or if the online platform only cares about the customers), non-extreme ratings are needed to approximate the customers' equilibrium payoff to the upper bound (Theorem 1). Interestingly, however, the system is seldom in such intermediary ratings. In other words, a mass concentration of data in extreme ratings arises endogenously from the system. Such "rating inflation" is empirically observed on popular rating systems.[4]

The optimal system for customers requires carefully designed transition rules to incentivize the expert to tell the truth whenever hired. Specifically, the higher the expert's discount factor is, or the more he cares about long-run interactions with customers, the less frequent and less dependent on the expert's choice of treatment the transitions between intermediary ratings must be. So, there is also a "rating persistence" pattern arising from the optimal system.

To maximize the expert's value from the interaction, we only need two ratings (Theorem 2). The optimal system is also a significant improvement in comparison to both the no-censoring and the full-censoring environments. In it, the expert is not tempted to reveal himself through an inadequate treatment choice and customers do not stop hiring him in the long run.

---

[2] Recently, Angi - former Angie's List - got caught up in a scandal concerning the aggregation of home service providers that advertise on the site in a "top-rate pros" category, even though such professionals were not highly rated among users (https://cbsn.ws/3GHRfDV).

[3] Our rating systems also capture a plausible account of the way platform users consume available data. Star ratings matter a great deal to online buyers. Recent surveys show that the star rating is the most or the second most important factor in online reviews that consumers pay attention to when judging a business or a product. See, for instance, surveys on Bright Local (2018, 2019). Customers in our model rely exclusively on the design of the rating system and the observation of the current rating to infer the type of expert they are hiring. Additionally, a reasonable assumption is that customers aggregate histories into common equivalent classes. Compte and Postlewaite (2015) motivate restricting players to play similarly across all histories in the same group of equivalence classes as a more realistic description of cooperative behavior in the long run.

[4] For example, eBay sellers have a median score of perfect rating (Nosko and Tadelis, 2015), and Airbnb has a percentage of properties with 4.5 stars or more (Zervas et al., 2021). Other rating systems, such as Amazon, Yelp, Uber and online labor marketplaces also exhibit this pattern. We stress there can be other explanations for this phenomenon, such as reporting biases; see, for example, Dellarocas and Wood (2008) and Filippas et al. (2018).

Both customers' and experts' equilibrium payoffs are constrained mainly because our class of information policies constrains the posterior belief spread. This happens because the restriction generates a strong connection between hiring and non-hiring ratings. At hiring ratings, the strategic expert must play a truth-telling strategy with at least some positive probability; otherwise, customers would not benefit from hiring. This means that he must generate the same signal of a bad expert. So the strategic expert visits non-hiring ratings sometimes in equilibrium. At non-hiring ratings, there is no additional evidence to separate types. Thus, the beliefs in such ratings account for the possibility of the expert being strategic, leading to an upper bound on the highest posterior belief about the expert being the bad type.

Our rating systems can generate enough data to support a trustful interaction between experts and customers, relative to no information and full information. Therefore, our paper has policy implications. Well-designed rating systems can alleviate the disarrangement between a patient's perception of a problem and a physician's diagnosis of it, for example; a problem discussed at the beginning of this introduction.

Related literature

We consider information censoring through finite memory as a device against bad reputational concerns. Ely and Välimäki (2003) construct a scoring mechanism that solves the bad reputation effect, but it depends on relaxing the assumption that customers are short-run. Being long-run, a customer has more incentives to hire the expert even if he sometimes provides the wrong treatment: by doing so today, the customer collects additional information about the expert's type to support better hiring decisions tomorrow. Mailath and Samuelson (2006) show that bad reputation is also avoidable when some customers are captive, that is, they choose to hire the expert regardless of his reputation. Our customers are short-lived and strategic.

Conceptually, the class of information policies we consider may resemble finite automata designed to generate some learning about an unknown state of the world. Thus we connect with the literature of learning with limited memory. We employ techniques from Hellman and Cover (1970) and Wilson (2014), both of which study optimal finite memory allocation when information is incomplete. As in Monte (2013, 2014), we apply such techniques in a reputation game. However, the design of the memory system comes from an intermediary in our model, whereas the cited papers discuss memory design from the perspective of the uninformed player. This means that all players in our model take the transition rules over memory states as given, but the actions in each memory state must be incentive compatible.

Lorecchio and Monte (2023) study rating systems to maximize the stationary probability of buyers buying a product of unknown quality. This paper, instead, has rating systems rating strategic service providers. As a consequence, reputational concerns generate endogenous information about the unknown state and additional incentive compatibility constraints. In other words, in Lorecchio and Monte (2023) the rating system rates a product based on exogenous signals about the product's quality, whereas here the system rates an expert based on signals that are endogenously generated by his behavior, which is itself dependent on the particular rating system.[5] Also, in this paper we focus more on consumer-optimal ratings.

Also related to our paper is Ekmekci (2011). In it, a rating system generates the right incentives for a long-run player to always take short-run players' most preferred action, to sustain a "good" reputation. Therefore, information censoring through finite memory is efficient in restoring good reputational concerns in repeated games with one-sided incomplete information.[6] In addition to discussing mechanisms to avoid bad reputational concerns, our system has some fundamental differences from Ekmekci's system. Most importantly, in his model, there is a permanent flow of informative signals about long-run players' type. In our model, whenever customers do not hire, there is no signal about his type. Technically, we deal with an interactive automaton.[7]

Other papers study memory restrictions to overcome the bad reputation effect. For instance, Sperisen (2018) discusses bounded recall, that is, when customers only remember treatment histories of a certain length. On the one hand, if the length is long enough, the unique long-run outcome for a patient expert is never to be hired. On the other hand, if the length is short enough, the expert has no incentives to provide the inadequate treatment, but customers' equilibrium payoff is similar to the one from a full-censoring environment (i.e., under zero recall). In fact, there is no "happy middle" under bounded recall, that is, it either tempts the expert into revealing himself in spite of harming customers through an inadequate treatment, or it provides the right incentives at the cost of not improving customers' equilibrium payoff, relative to the trivial information censoring.

Although we focus on finite rating sets, our setting can accommodate countably infinite sets as well. In particular, the rating set can be the entire space of infinite sequences of public outcomes and the transition rule can reproduce the full-

[5] Moreover, in the bad reputation environment we consider, it is more intricate to induce information generation. If the rating system seeks to generate only a "good" signal from the expert in hiring ratings, it may require him to provide inadequate treatment to customers. Not only this has bad payoff consequences for both parties, but it may also result in customers not finding optimal to hire the expert in the first place, contradicting the behavior that the system would like to induce on the customers.

[6] Fudenberg and Levine (1989, 1992) have shown that if there is a small common knowledge probability that the long-run player is committed to the short-run players' most preferred action, it can be sequentially rational for this long-run player to always play such preferred action. Cripps et al. (2004) however proved that reputational concerns are insufficient to curb incentives for the long-run player to deviate when his actions are imperfectly monitored.

[7] Vong (2021) studies rating systems in a reputation game similar to Ekmekci (2011), but with a different commitment type of long-lived player. In it, the intermediary designs a certification scheme to know the type of the long-lived player prior to the construction of the rating system. Instead, our designer only observes the outcomes of the interactions with short-run players, which are signals about the expert's type.

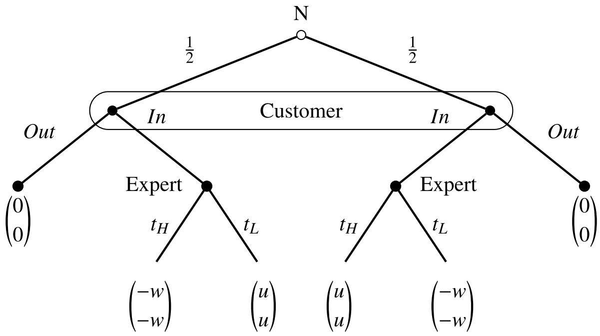

Fig. 1. The stage game. Nature draws a problem H or L with equal probability. A customer chooses whether to hire the expert or not. If hired, expert observes the problem and proposes treatment t_H or t_L. A right treatment generates u for both players; a wrong one generates $-w$.

information environment from Ely and Välimäki (2003) as well as the bounded-recall environment from Sperisen (2018).[8] Most importantly, our model allows for stochastic transitions between ratings. By doing so, we show that, even with two ratings, for a range of prior beliefs, we can construct a system that generates the right incentives for the expert and it is informative enough to improve customers' equilibrium payoff relative to the full-censoring environment.[9,10]

Our system is time-independent and we focus on stationary strategies, since we assume that customers are unaware of calendar time. We view this as a design of a simple information rule.[11] Simpler rules have advantages over more complicated ones. Dynamic information policies can depend on public histories in a quite complex fashion as well as require customers to keep track of all possible outcomes of past interactions. As in Compte and Postlewaite (2015), we view our rating systems as a more plausible description of how customers reason about the expert's type upon observing a given rating.

Pei (2022) studies sustainable reputations under observational restrictions that are similar to ours, despite the payoff structure he considers being different.[12] Specifically, he assumes that short-lived players only see "summary statistics" about the past actions of the long-lived player, or unordered, time-bounded observations of the past play. In his model, he shows that short-run players can approximately attain their first-best in all Nash equilibria, provided that memory is short enough. In our model, we show that short-lived players attain approximately their first-best in a system with a sufficiently long memory.

2. Model

The setting is an infinitely repeated game between one long-run player (the expert) who interacts with a sequence of identical myopic players (the customers). Time is indexed by $t = 1, 2, 3, \ldots$, and at each period t a different customer interacts with the long-run player. We reproduce the underlying model of Ely and Välimäki (2003), but we restrict strategies to be functions of ratings and not on finer details of the public history.

At the beginning of every stage game, Nature draws a problem $\theta \in \{H, L\}$, each happening with equal probability. Without observing θ, a customer decides first whether to hire the expert (In) or not (Out). If she does not hire, both players get an outside option payoff which is normalized to zero. If she hires, the expert then perfectly observes θ and decides the level of service to provide. An appropriate treatment for H would be t_H and an appropriate treatment for L would be t_L. The customer observes the treatment and the payoffs are realized: the right treatment generates a payoff of u and the wrong one generates a payoff of $-w$. The stage game is summarized in Fig. 1.

There is incomplete information regarding the expert's type. Specifically, with common knowledge probability $\rho \in (0, 1)$, the expert is a type committed to a pure strategy: he always provides the high-cost treatment t_H. We denote this behavioral

[8] We formally demonstrate how a rating system with a countably infinite rating set can reproduce the unbounded and bounded-recall environments in the Online appendix.

[9] Sperisen (2018) also considers bad reputation under fading recall, that is, when customers' memory decays exponentially over time. Bad reputation is avoidable provided that memory decays quickly enough. An exponentially decaying memory requires a computationally complex automaton, with infinite states, each one with a probability of being directly reached from previous states. Our optimal systems are finite and do not require customers to keep track of past sequences of public outcomes.

[10] Under a different information technology, but dealing with a related problem, Lillethun (2017) studies dynamic information disclosure in repeated games with reputational concerns and considers the bad reputation game as an application.

[11] We show in the Online appendix that calendar-time awareness is not the driver of our main results in this bad reputation environment. That is, under finite rating systems, there is still an upper bound on induced beliefs even if players know their place in the queue.

[12] He focuses on results about sustainable reputations in stage games with supermodular payoffs, which our model lacks. His payoff structure also implies that short-lived players observe a persistent flow of information about the long-run players' type.

"bad" type by B. With probability (1-ρ), he is a strategic type,[13] denoted by S, and his payoffs at the end of each period are identical to the customers' payoffs.[14]

Assume $w > u > 0$. It is straightforward to see that the expert strictly prefers to provide the correct treatment when hired in the stage game in absence of a repeated setting. Let $\nu(\rho)$ denote a customer's expected payoff from hiring in the stage game as a function of the initial prior. This is given by

$$\nu(\rho) = \rho\left(\frac{u-w}{2}\right) + (1-\rho)u.$$

The non-hiring payoff is zero to both players, so she will not hire if $\nu(\rho) < 0$. A simple manipulation of the above equation shows that this is equivalent to ρ being higher than a belief threshold $\rho^* \in (0,1)$, defined below. Relying only on the initial prior, and given that customers are myopic, if $\rho \leq \rho^*$, they will hire the expert; if $\rho > \rho^*$, they will not do so.

$$\rho^* = \frac{2u}{u+w}.$$

2.1. Bad reputation

Let us now discuss the bad reputation result in Ely and Välimäki (2003). Assume first that $\rho > \rho^*$. Since the myopic customers do not care about the information externalities they may bring to future customers by hiring the expert, they never hire and the expert will never get the chance to reveal himself. Thus, his average discounted payoff in equilibrium is zero.

Assume now that $\rho \leq \rho^*$. Then, in principle, hiring could take place initially. Nevertheless, Ely and Välimäki show that if customers get to see all past treatments chosen by the expert (but not the associated problems) and choose according to these public histories, the expert will rarely be hired in any (Nash) equilibrium if he is sufficiently patient (or, if he cares about interacting with customers infinitely often). Moreover, if incentives to play the wrong treatment to avoid this zero payoff equilibrium are strong enough, he will never be hired. History-dependent strategies and reputation effects harm both the expert and the customers in this game.

Here is the intuition for this result. At every period, hiring is possible only if the expert is playing correct actions with positive probability for every problem, including treatment t_H for problem H. This implies that there is always a chance of the path of play reaching a critical history, after which one more observation of the high-cost treatment t_H leads to prohibitively high posterior belief about the expert being bad. At such history, either he eventually proposes t_H and never gets hired again, or the temptation to overchoose the low-cost treatment t_L is too strong not to be noticed by future customers. Anticipating the expert's incentives, they will refuse to hire him and risk being treated incorrectly. If the expert discounts future payoffs at a rate δ, he can expect to earn the zero-payoff on average if he is sufficiently patient.

2.2. Rating systems

We now assume that all customers have access to only one source of information - a rating system. Specifically, a rating system $\mathcal{M} := (M, \varphi, \varphi_0)$ consists of (i) a finite set of ratings M; (ii) a transition function φ determining a probability distribution over ratings at period t as a function of the rating at period $t-1$ and the outcome of the customer's interaction with the expert in that period (t_H or t_L in case she hired him; Out otherwise); and (iii) an initial probability distribution φ_0 over ratings. We assume that the designer has commitment power to design the rating system.

Customers can only condition their behavior on ratings, not on finer details of the public history. Thus, we assume that their strategies can not depend on calendar time. They take the system as given and their strategies are a map from the current rating to a decision to hire or not, as follows[15]:

$$\alpha : M \to \{0, 1\}.$$

We will say hat i is a hiring rating if $\alpha = 1$ and a non-hiring rating if $\alpha = 0$.

The expert observes the realized rating at every period. Since this is the only information customers have, we will focus on equilibria in which the expert's choice of treatment will vary only across ratings and problems, not across time. Thus, the expert's strategies are maps from current ratings and current problems to a probability of providing the correct treatment, as follows:

---

[13] Unless stated otherwise, when we refer to the expert, we will be referring to this strategic type.

[14] We follow Ely and Välimäki's benchmark setting of perfectly aligned payoffs because it highlights the striking effects of bad reputational concerns under full memory and neatly pinpoints key aspects of bad reputation games. In their paper, they also analyze the case in which the bad type is a strategic type biased in favor of the high-cost treatment, but not informed about each customer's problem. Although their analysis to this case is more involved, they show that bad outcomes from reputational concerns still arise.

[15] It is without loss of generality to focus on deterministic strategies, since we can replicate the effects of a random hiring decision with a proper redesign of transition probabilities. We formally show this in the Online appendix.

$$\beta^H : M \to [0,1],$$
$$\beta^L : M \to [0,1].$$

We assume that customers compute the probability distribution over the public histories as if the game had been going on for a long time. Thus, they compute beliefs using steady-state probabilities (or time-average convergence). To understand how these probabilities arise, note that any given rating system $\mathcal{M}$ and strategy profile $(\alpha, \beta^H, \beta^L)$ define Markov matrices T^S and T^B for the strategic and the bad expert, respectively. To simplify notation, let $\varphi^y_{mm'}$ represent the probability of transitioning from rating m to m' upon the observation of $y = H$ (corresponding to t_H), $y = L$ (corresponding to t_L) or $y = Out$. Then the entries of T^S and T^B are

$$\tau^S_{mm'} = \alpha_m[\gamma_m \varphi^H_{mm'} + (1-\gamma_m)\varphi^L_{mm'}] + (1-\alpha_m)\varphi^{Out}_{mm'}, \tag{1}$$
$$\tau^B_{mm'} = \alpha_m \varphi^H_{mm'} + (1-\alpha_m)\varphi^{Out}_{mm'}, \tag{2}$$

where $\gamma_m := \frac{1}{2}\beta^H_m + \frac{1}{2}(1-\beta^L_m)$ represents the probability of observing treatment t_H at m and $1-\gamma_m$ the probability of observing treatment t_L, if the expert is strategic. For now, we will restrict the analysis to Markov matrices that are irreducible for the bad expert,[16] but in section 5 we discuss the general case. From a well-known result in Markov processes,[17] there will be invariant distributions $f^B := (f^B_m)_{m\in M}$ and $f^S := (f^S_m)_{m\in M}$. Moreover, in each distribution, all probabilities will be positive (Lemma 2 in the Appendix). Customers use such distributions to compute updated beliefs about the expert being bad at every rating reached with positive probability, as defined below.

$$\rho_m = \frac{\rho f^B_m}{\rho f^B_m + (1-\rho) f^S_m}. \tag{3}$$

Define $\beta_m := \frac{1}{2}\beta^H_m + \frac{1}{2}\beta^L_m$ as the expert's probability of telling the truth at rating m. Customers' expected payoff from hiring is

$$\nu(\rho_m) = \rho_m\left(\frac{u-w}{2}\right) + (1-\rho_m)[\beta_m u - (1-\beta_m)w]. \tag{4}$$

Customers' strategy must be incentive compatible, that is, whenever they are willing to hire, the expected payoff from doing so must be equal to or greater than zero (the payoff from not hiring). Thus, customers find optimal to hire for every m reached with positive probability in the long-run such that $\nu(\rho_m) \geq 0$.

Let V^θ_m denote expert's continuation value from a rating system and a strategy profile, at rating m, after being hired and after observing problem $\theta \in \Theta := \{H, L\}$. Let $V_m := \frac{1}{2}V^H_m + \frac{1}{2}V^L_m$. Then, the continuation value is given by

$$V^\theta_m := (1-\delta)[\beta^\theta_m u - (1-\beta^\theta_m)w] + \delta \sum_{m'\in M}[\beta^\theta_m \varphi^{y=\theta}_{mm'} + (1-\beta^\theta_m)\varphi^{y\in\Theta\setminus\{\theta\}}_{mm'}]V_{m'}. \tag{5}$$

The expert's strategy must be optimal, that is, whenever hired at rating m and upon observing problem θ, β^θ_m must maximize the continuation value defined above.

For any given rating system $\mathcal{M}$ and parameters (ρ, δ), we define an equilibrium as a strategy profile and a posterior belief distribution $(\rho_m)_{m\in M}$ such that (i) customers are taking optimal actions given their posterior beliefs; (ii) the expert is taking optimal actions whenever hired and (iii) posterior beliefs are consistent with Bayes rationality and the distributions f^B and f^S. We state the concept below.

Definition 1. An equilibrium is a strategy profile $(\alpha, \beta^H, \beta^L)$ as well as a posterior belief distribution $(\rho_m)_{m\in M}$ such that, for every $m \in M$,

1. ρ_m is consistent (derived from Bayes' rule and the stationary distributions f^B and f^S, as in equation (3));
2. $\nu(\rho_m) \geq 0 \Rightarrow \alpha_m = 1$ and $\nu(\rho_m) < 0 \Rightarrow \alpha_m = 0$, where $\nu(\rho_m)$ is defined in equation (4);
3. for $\theta \in \{H, L\}$, β^θ_m maximizes V^θ_m whenever $\alpha_m = 1$, as in equation (5).

To understand the mechanics of our model and how to compute beliefs, as well as to illustrate how simple rating systems help overcome market failures and improve upon both the full-memory and the no-memory cases, consider the following example.

[16] A Markov Matrix is irreducible if for any pair of ratings (i, j), there is a positive probability of moving from i to j in finite time. See Section 5.
[17] See for instance Stokey and Lucas (1989), Theorem 11.2.

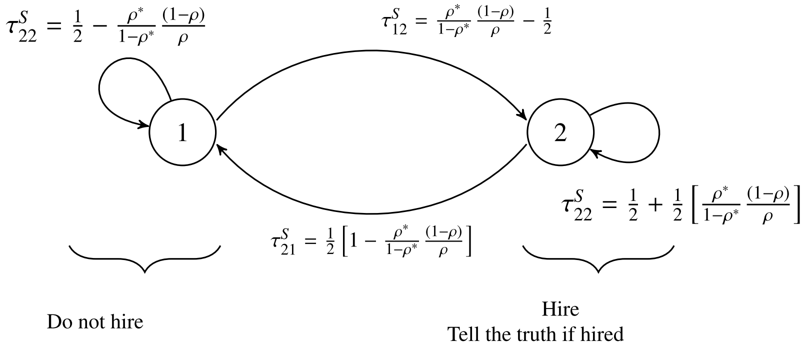

Fig. 2. A binary and stochastic rating system in which the expert tells the truth if hired and customers hire only in rating 2. With probability $\varphi^H_{21} = 1 - \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho}$, the expert moves from the hiring rating to the non-hiring one after telling the truth at problem H. The observation of treatment t_L never moves the system from 2 to 1. With probability $\varphi^{Out}_{12} = \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} - \frac{1}{2}$, the system moves from the non-hiring rating to the hiring one.

Example 1. The rating set is binary and the transition rule is stochastic. The customers hire in rating 2, but not in rating 1. Starting at 2, every time the expert provides treatment t_H, the rating moves to 1 with probability $\varphi^H_{21} = 1 - \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho}$; otherwise, the system remains in 2. Once in 1, the system moves to 2 with probability $\varphi^{Out}_{12} = \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} - \frac{1}{2}$. Assume that $\frac{2\rho^*}{1+\rho^*} > \rho > \rho^*$ and note that under these parameters the transition probabilities are well-defined. Our equilibrium candidate is (i) $\alpha_1 = 0$ and $\alpha_2 = 1$; (ii) $\beta^H_2 = \beta^L_2 = 1$. This system and the strategies that are part of an equilibrium candidate are represented in Fig. 2 above.[18]

For such a system, recalling that the entries of the transition matrices T^S and T^B are given by equations (1) and (2) respectively, we get

$$T^B = \begin{array}{c} \\ 1 \\ 2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} \quad 1 \qquad\qquad & \qquad 2 \end{array} \\ \begin{pmatrix} \frac{1}{2} - \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} & \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} - \frac{1}{2} \\ 1 - \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} & \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} \end{pmatrix} \end{array} \qquad T^S = \begin{array}{c} \\ 1 \\ 2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} \quad 1 \qquad\qquad & \qquad 2 \end{array} \\ \begin{pmatrix} \frac{1}{2} - \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} & \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} - \frac{1}{2} \\ \frac{1}{2}\left[1 - \frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho}\right] & \frac{1}{2} + \frac{1}{2}\left[\frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho}\right] \end{pmatrix} \end{array}$$

Both matrices do not have absorbing subsets, so we can compute the stationary distributions using the identities $f^B \cdot T^B = f^B$ and $f^S \cdot T^S = f^S$. This leads to the distributions

$$f^B_1 = \frac{2(\rho - \rho^*)}{(1-\rho^*)\rho} \qquad f^B_2 = \frac{\rho^*(1-\rho) - (\rho - \rho^*)}{(1-\rho^*)\rho},$$
$$f^S_1 = \frac{\rho - \rho^*}{\rho^*(1-\rho)} \qquad f^S_2 = \frac{\rho^*(1-\rho) - (\rho - \rho^*)}{\rho^*(1-\rho)}.$$

From equation (3), the equilibrium beliefs will be

$$\rho_1 = \frac{2\rho^*}{2+\rho^*} \qquad \rho_2 = \rho^*.$$

It remains to check whether the strategy profile and the beliefs form an equilibrium (Definition 1). Note that $\nu(\rho_1) < 0$ since $\rho_1 > \rho^*$, so it is optimal for customers not to hire in rating 1 ($\alpha_1 = 0$); likewise, $\nu(\rho_2) = 0$ since $\rho_2 = \rho^*$, so it is a best response for customers to hire in rating 2 ($\alpha_2 = 1$). Therefore, the customers' strategies are incentive compatible.

It is clearly optimal for the expert to choose t_L upon the observation of problem L in rating 2: this maximizes his expected payoff whenever hired and upon observing problem L, as well as the likelihood of remaining in the hiring rating 2.

It is not immediate to see that it is optimal for the expert to provide adequate treatment after observing problem H: recommending t_L instead reduces his expected payoff, but avoids a downgrade to the non-hiring rating, which is costly in terms of continuation payoffs. To prove the optimality of providing the correct treatment, first recall that, from equation (5), the expert's continuation value in the hiring rating 2 and upon the observation of problem $\theta = H$, given any truth-telling probability $\beta^H_2 \in [0, 1]$, is

$$V^H_2 := (1-\delta)[\beta^H_2 u - (1-\beta^H_2)w] + \delta \sum_{m \in M} [\beta^H_2 \varphi^H_{2m} + (1-\beta^H_2)\varphi^L_{2m}] V_m.$$

[18] We also set $\beta^H_1 = \beta^L_1 = 1$, although in our candidate equilibrium the expert will not be hired in rating 1.

Conditional on any rating system $\mathcal{M} := (M, \varphi, \varphi_0)$ and any customers' strategy α, the expert is solving a dynamic programming problem with bounded per-period payoffs that are discounted at rate $0 < \delta < 1$. A stationary (conditional) strategy β^H is optimal if and only if β_i^H attains the maximum value of equation (5), for each $i \in M$.[19] In particular, $\beta_2^H = 1$ is optimal given the rating system and the customers' strategy of this example if

$$(1-\delta)u + \delta[\varphi_{21}^H V_1 + \varphi_{22}^H V_2] \geq \beta_2^H\{(1-\delta)u + \delta[\varphi_{21}^H V_1 + \varphi_{22}^H V_2]\} + (1-\beta_2^H)\{(1-\delta)(-w) + \delta V_2\},$$

which simplifies to:

$$(1-\delta)u + \delta[\varphi_{21}^H V_1 + \varphi_{22}^H V_2] \geq (1-\delta)(-w) + \delta V_2,$$

or, even more concisely, $(1-\delta)(u+w) \geq \delta\varphi_{21}^H[V_2 - V_1]$. To see this is indeed true, note that, from the continuation value in rating 2,

$$V_2 = (1-\delta)u + \delta\left[\tau_{21}^S V_1 + \tau_{22}^S V_2\right] \Rightarrow V_2 - V_1 = \frac{(1-\delta)[u - V_1]}{1-\delta+\delta\tau_{21}^S}.$$

Thus,

$$\begin{aligned}
\delta\varphi_{21}^H[V_2 - V_1] &= (1-\delta)[u - V_1]\left\{\frac{\delta\varphi_{21}^H}{1-\delta+\delta\tau_{21}^S}\right\}, \\
&< (1-\delta)[u+w]\left\{\frac{\delta\frac{1}{2}\varphi_{21}^H}{1-\delta+\delta\tau_{21}^S}\right\}, \\
&< (1-\delta)(u+w).
\end{aligned}$$

The first inequality follows from $V_1 > 0 > \frac{u-w}{2}$ (see the expression for V_1 below); the second inequality follows from $\tau_{21}^S = \frac{1}{2}\varphi_{21}^H$. Therefore, for this system and for any prior $\rho^* < \rho < \frac{2\rho^*}{1+\rho^*}$, the strategy profile and posterior beliefs form an equilibrium, because (i) customers optimally hire in rating 2 but not in rating 1; (ii) truth-telling is optimal in rating 2 and (iii) beliefs are consistent.

Note that, in this equilibrium, at least for a range of prior values, reputational concerns do not tempt the expert to play the wrong treatment to reveal himself. Note as well that the expert's equilibrium payoff is bounded away from zero, even as he becomes increasingly patient. Indeed, solving for the expert's continuation values, we get

$$V_1 = \left\{\frac{\delta\left[\frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} - \frac{1}{2}\right]}{1-\delta+\frac{\delta}{2}\left[\frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho}\right]}\right\}u \qquad V_2 = \left\{\frac{1-\delta+\delta\left[\frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho} - \frac{1}{2}\right]}{1-\delta+\frac{\delta}{2}\left[\frac{\rho^*}{1-\rho^*}\frac{(1-\rho)}{\rho}\right]}\right\}u.$$

Because the system starts in rating 2, the expert's equilibrium payoff is V_2. For $\delta \to 1$, the equilibrium payoff converges to

$$f_2^S u = \left[\frac{\rho^*(1-\rho) - (\rho-\rho^*)}{\rho^*(1-\rho)}\right]u > 0.$$

Finally, note that the expert's equilibrium payoff is also higher than both the payoff under full censoring, since in that case and for $\rho > \rho^*$, no customer would ever hire the expert.[20]

3. Optimal rating systems for customers

We begin by examining optimal rating systems for customers. They maximize the expected value of the customers' payoffs at each rating with respect to the probability of reaching such ratings, provided that the induced strategy profile is an equilibrium. Specifically, they maximize the following value from the interaction:

$$v = \sum_{m\in M} f_m \alpha_m \nu(\rho_m), \tag{6}$$

where $f_m := \rho f_m^B + (1-\rho)f_m^S$. We first prove that, in any rating system inducing at least one hiring rating and at least one non-hiring rating, there is an upper bound on induced beliefs. We then prove that this leads to an upper bound on

[19] See for instance Bertsekas (2015), Proposition 1.2.3.

[20] In fact, we will later prove in Section 4 that this illustrative rating system and equilibrium actually maximizes the expert's value from the interaction for the specific case of when the expert evaluates his payoffs according the stationary distribution.

the customers' value from the interaction. Given this constraint, we construct a rating system that gets arbitrarily close to customers' maximum equilibrium payoff, for every prior belief below the derived belief upper bound.

Abusing notation, we let $|M| = M$ and without loss of generality we label ratings such that $\rho_1 \geq \ldots \geq \rho_M$. Note that this means that the long-run relative frequencies have a non-increasing order, that is,

$$\frac{f_1^B}{f_1^S} \geq \frac{f_2^B}{f_2^S} \geq \ldots \geq \frac{f_M^B}{f_M^S}.$$

Before proceeding, we need a definition and a notation. First, we say that a rating system induces an informative equilibrium if it induces an equilibrium such that at least one rating is a hiring rating and at least one rating is a non-hiring rating.

Definition 2. A rating system induces an informative equilibrium if it induces an equilibrium such that at least one rating is a hiring rating and at least one rating is a non-hiring rating.

Second, let us define a useful statistic:

$$\lambda = \frac{\rho^*}{(1-\rho^*)} \frac{(1-\rho)}{\rho}.$$

If $\rho \leq \rho^*$, we have $\lambda \geq 1$; if $\rho > \rho^*$, then $\lambda < 1$. This represents the prior bias in favor of hiring the expert. Define as well $\mathcal{O}$ as the set of non-hiring ratings (*Out*) and $\mathcal{I}$ as the set of hiring ratings (*In*). Again from equation (4), we can rewrite customers' incentive-compatibility inequalities as

$$\frac{f_i^B}{f_i^S} > \lambda \left[\beta_i - (1-\beta_i)\frac{w}{u}\right] \;\; \forall\, i \in \mathcal{O}; \tag{7}$$

$$\frac{f_i^B}{f_i^S} \leq \lambda \left[\beta_i - (1-\beta_i)\frac{w}{u}\right] \;\; \forall\, i \in \mathcal{I}. \tag{8}$$

Let us present our first result. Proposition 1 below, the proof of which is in the Appendix, proves that there is a third restriction for equilibria with rating systems inducing an informative equilibrium. We follow a similar logic as in Lorecchio and Monte (2023). The main difference between that paper and this one is that here a strategic agent (the expert) is being rated, so the induced stationary probability depends on the rating system but also on the expert's strategy, which, in turn, is a best response to the strategy of the customers and the rating system. In that paper, the rating system induced a type-dependent stationary probability, and the signals were generated whenever the product was bought. In Proposition 1, we adapt the proof of a proposition in that paper to our (strategic) environment.

Proposition 1. *In any rating system inducing an informative equilibrium, all non-hiring ratings i obey*

$$\frac{f_i^B}{f_i^S} \leq 2\lambda.$$

Proposition 1 implies that in all non-hiring ratings, customers will hold induced beliefs about a bad type of expert no higher than $\bar{\rho}$, defined by

$$\bar{\rho} = \frac{2\rho^*}{1+\rho^*}.$$

If $\rho \geq \bar{\rho}$, there can be no rating system inducing an informative equilibrium. This follows from (i) in equilibrium, the *ex-ante* expected belief is equal to the prior - or $\sum_{i=1}^{M} f_i \rho_i = \rho$; (ii) every non-hiring rating in an informative equilibrium must induce a belief no higher than $\bar{\rho}$. Therefore, to support an informative equilibrium with $\rho \geq \bar{\rho}$, all hiring ratings would have to satisfy $\sum_{i \in \mathcal{I}} f_i(\rho_i - \bar{\rho}) \geq 0$. This means that the system would have to exhibit a hiring rating with a belief at least $\bar{\rho}$. But then it is not optimal for customers to hire under such a belief. For this reason, from now, we will impose the following assumption.

Assumption 1. The prior belief about expert being bad is such that $\rho < \bar{\rho}$.

From the point of view of the customers, it would be desirable that they learned the expert's type. More specifically, they would benefit from knowing whether they are hiring the strategic expert and not hiring the bad one. However, from Proposition 1, we know that the highest belief about the expert being bad we can possibly achieve is $\bar{\rho}$. From the irreducibility assumption, we also know that the lowest belief about the expert being bad we can possibly achieve is higher than zero. This creates an upper bound on customers' maximum equilibrium payoff. We characterize such bound in the next proposition.

Proposition 2. *In any rating system inducing an informative equilibrium, for every prior belief $\rho \in (0, \bar{\rho})$, customers' maximum equilibrium payoff is bounded above:*

$$v \leq u\left[\frac{\bar{\rho}-\rho}{\bar{\rho}}\right].$$

Proof. As before, suppose non-hiring ratings belong to the set $\mathcal{O}$ and hiring ones to $\mathcal{I}$. Whenever customers do not hire, they get a zero payoff. Whenever they hire in $i \in \mathcal{I}$, they get payoff $\nu(\rho_i)$, as given by equation (6). Therefore, for a rating system inducing an informative equilibrium,

$$\begin{aligned} v &= \sum_{i\in\mathcal{I}} f_i \nu(\rho_i), \\ &= \sum_{i\in\mathcal{I}} f_i \left\{\rho_i\left(\frac{u-w}{2}\right) + (1-\rho_i)\left[\beta_i u - (1-\beta_i)w\right]\right\}, \\ &\leq \sum_{i\in\mathcal{I}} f_i \left\{\rho_i\left(\frac{u-w}{2}\right) + (1-\rho_i)u\right\}. \end{aligned}$$

The inequality follows because, in any $i \in \mathcal{I}$, $\beta_i = 1$ maximizes customers' hiring payoff. Define $f_{\mathcal{I}} := \sum_{i\in\mathcal{I}} f_i$ and $\rho_{\mathcal{I}} := \frac{\sum_{i\in\mathcal{I}} f_i\rho_i}{f_{\mathcal{I}}}$. Then the above bound becomes

$$v \leq f_{\mathcal{I}}\left\{\rho_{\mathcal{I}}\left(\frac{u-w}{2}\right) + (1-\rho_{\mathcal{I}})u\right\}. \tag{9}$$

In equilibrium, it must be that $\sum_{i=1}^{M} f_i\rho_i = \rho$. Thus, if we partition the rating set into non-hiring ratings and hiring ones, we obtain

$$\sum_{i\in\mathcal{I}} f_i\rho_i + \sum_{i\in\mathcal{O}} f_i\rho_i = \rho.$$

Define as well $\rho_{\mathcal{O}} := \frac{\sum_{i\in\mathcal{O}} f_i\rho_i}{1-f_{\mathcal{I}}}$. From Proposition 1, $\rho_{\mathcal{O}} \leq \bar{\rho}$. Rewriting the above equation in terms of $f_{\mathcal{I}}$, $\rho_{\mathcal{I}}$, $\rho_{\mathcal{O}}$ and using the result that $\rho_{\mathcal{O}} \leq \bar{\rho}$, we get an upper bound on $f_{\mathcal{I}}$:

$$\rho = f_{\mathcal{I}}\rho_{\mathcal{I}} + (1-f_{\mathcal{I}})\rho_{\mathcal{O}} \leq f_{\mathcal{I}}\rho_{\mathcal{I}} + (1-f_{\mathcal{I}})\bar{\rho} \;\Rightarrow\; f_{\mathcal{I}} \leq \frac{\bar{\rho}-\rho}{\bar{\rho}-\rho_{\mathcal{I}}}. \tag{10}$$

Substituting equation (10) into equation (9) leads to

$$v \leq (\bar{\rho}-\rho)\left\{\frac{\rho_{\mathcal{I}}\left(\frac{u-w}{2}\right) + (1-\rho_{\mathcal{I}})u}{\bar{\rho}-\rho_{\mathcal{I}}}\right\}.$$

The right-hand side of the above inequality is decreasing in $\rho_{\mathcal{I}}$, provided that $\bar{\rho} > \frac{2u}{u+w} := \rho^*$. This is indeed satisfied, since $\bar{\rho} = \frac{2\rho^*}{1+\rho^*}$ and $\rho^* \in (0,1)$. Thus, the highest value is $u\left[\frac{\bar{\rho}-\rho}{\bar{\rho}}\right]$. □

The upper bound from Proposition 2 is best seen in Fig. 3 below. The solid, black line represents the payoff under the one-shot interaction and the loosely dotted, red line the bound derived in the previous proposition. The red, outer dashed line represents the payoff customers would get if the system could perfectly separate types (and have the strategic expert telling the truth whenever hired).

For every positive prior belief satisfying Assumption 1 and any discount factor between zero and one, we can construct a sequence of rating systems inducing an informative equilibrium in which customers' payoff approaches the maximum bound derived in Proposition 2. Along this sequence of systems, customers hire in every rating except rating 1 and expert tells the truth whenever hired. This is true even for discount factors arbitrarily close to one and hence under a possibly strong temptation to manipulate reputation by providing t_L for problem H. The construction of such a sequence of systems is given in the proof of the next theorem, which is in the Appendix, but we provide a sketch here.

In any rating system inducing an informative equilibrium, at least one rating must have a belief (weakly) lower than the prior, and at least one rating must have a belief (weakly) higher than the prior. It suffices to induce non-hiring in rating 1 only and hiring in the other ratings. The intermediary hiring ratings will be useful to learn more about the expert's type.

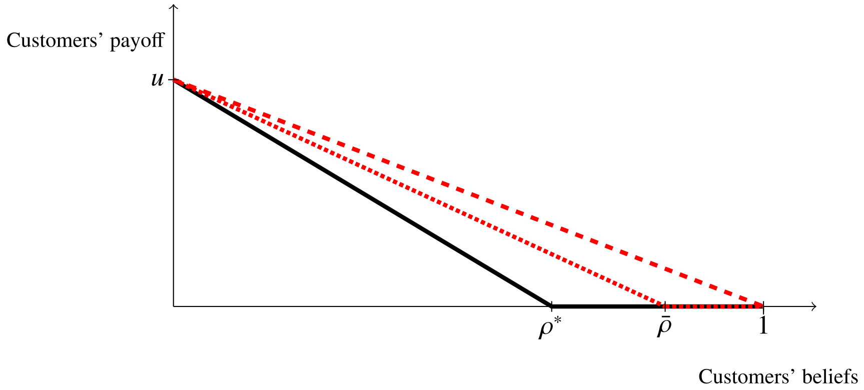

Fig. 3. Customers' equilibrium payoffs. The solid black line represents the one-shot payoff. The dashed, red line represents the payoffs under a perfect separation of types and the dotted, red line represents an upper bound on the value customers can get from the interaction with an expert through a rating system. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Since the optimal payoff for the customers is a weighted average of the payoffs in each rating and since the weights must be such that $\sum_{i=1}^{M} f_i \rho_i = \rho$, then, ideally, the system would induce the highest posterior belief about expert being bad (ρ_1) equal to the belief upper bound $\bar{\rho}$. For every prior belief $\rho \in (0, \bar{\rho})$, the system must also induce at least one posterior belief lower than the prior. To do so, for any $2 \le M < \infty$, we first define a belief lower bound:

$$\underline{\rho}(M) := \frac{\rho^* \left(\frac{1-\rho^*}{2}\right)^{M-2}}{1-\rho^* + \rho^* \left(\frac{1-\rho^*}{2}\right)^{M-2}}.$$

Note that $\underline{\rho}(M) \le \rho^*$, with strict inequality for $M > 2$. Note as well that, for any prior belief satisfying Assumption 1, we can find a value of M_0 high enough so that $\rho > \underline{\rho}(M_0)$. Therefore, for any rating system with M_0 ratings, if we can induce beliefs ρ_1 and ρ_{M_0} such that $\rho < \rho_1 \le \bar{\rho}$ and $\underline{\rho}(\overline{M}_0) \le \rho_{M_0} < \rho$, then an informative equilibrium is at least possible.

Second, fixing the number of ratings in the system at M_0, we construct a sequence of irreducible systems and informative equilibria along which the expert gets hired in every rating except rating 1. Moreover, the sequence of induced beliefs has ρ_1 approaching $\bar{\rho}$ and ρ_{M_0} approaching $\underline{\rho}(M_0)$ under a truth-telling strategy. To do so, we carefully choose the sequence of transition probabilities so that, at intermediary ratings, upgrades are more likely than downgrades after t_L and downgrades are more likely than upgrades after t_H.

We also carefully choose the transition probabilities out of extreme ratings to be close to zero, while keeping posterior beliefs in intermediary ratings well-defined. Specifically, at rating 1, for every element $t \in \mathbb{N}$ in the sequence, there is an exit probability τ_t to rating 2. At rating M_0, there exists a downgrade probability κ_t to $M_0 - 1$ if the high-cost treatment is observed. Under a low-cost treatment at M_0, the system stays there. With a proper choice of $\frac{\tau_t}{\kappa_t}$ and the intermediary transitions, we can have $\rho_1 \to \bar{\rho}$ and $\rho_{M_0} \to \underline{\rho}(M_0)$ even if $\tau_t \to 0$ and $\kappa_t \to 0$. In words, as the system spends less time in intermediary ratings, the belief spread approaches the highest possible value for the given construction.

Third, we can set up posterior beliefs in a way that is incentive compatible for customers to hire in every rating except 1, but we must adjust transitions so that the expert finds it optimal to tell the truth whenever hired, for every discount factor. This is tricky: the more patient the expert is, the more he cares about being caught up in the extreme rating in which customers hire. He then has incentives to game the system by providing cheap treatment to a severe problem.

To account for this, not only we adapt the intermediary transitions in a way that the expert is downgraded with the same probability as he is upgraded if he tells the truth, but we also make the transitions dependent on the discount factor in a way that, higher the discount, the lower the chances of leaving intermediary ratings, no matter the proposed treatment. The intuition for this is simple: since with no memory there is no conflict of interest (the expert and customers' utilities coincide), it suffices to make intermediary ratings relatively more persistent for a higher discount.

Finally, to have customers' equilibrium payoff arbitrarily close to the upper bound from Proposition 2, it is just a matter of choosing the number M of ratings in the system, provided that is at least M_0, depending on how close to zero we want $\underline{\rho}(M)$ to be, and then setting up the transition probabilities according the previous paragraphs. Intuitively, higher the number of ratings in this construction, lower the chances of a bad expert reaching the extreme hiring rating, since upgrades are more likely under the provision of the low-cost treatment.

Theorem 1. *For any prior belief $\rho \in (0, \bar{\rho})$, any discount factor $\delta \in (0, 1)$, and $\forall \varepsilon > 0$ there is a sufficiently large rating set M and a sequence of rating systems $\mathcal{M}_t := (M, \varphi_t, \varphi_0)$ and a number $T > 1$ such that (i) each element of the sequence induces an informative equilibrium; (ii) the customers' payoff approaches the upper bound derived in Proposition 2: $v_t > u\left[\frac{\bar{\rho}-\rho}{\bar{\rho}}\right] - \varepsilon, \forall t > T$. This sequence of rating systems has customers hiring in every rating except rating 1 and the expert telling the truth whenever hired.*

In constructing such sequence of rating systems, we note that, no matter how patient the expert is, he has no temptation to manipulate reputation and there is always a positive chance of him being hired infinitely often. Therefore, even for a sequence of discount factors approaching one, the expert's equilibrium payoff along such sequence is bounded away from zero. This is fundamentally different from what the results in Ely and Välimäki (2003) imply under no-information censoring. To highlight such difference, we state this finding as Corollary 1 below.

Corollary 1. *Given any prior belief $\rho \in (0, \bar{\rho})$, for any sequence of discount factors approaching one, there is a sequence rating systems satisfying the properties described in Theorem 1 and leading to a sequence of expert's equilibrium payoffs uniformly bounded away from zero.*

4. Optimal rating systems for the expert

We now examine optimal rating systems for the expert. Note that if $\rho \leq \rho^*$ any customer would hire under the prior belief. In this case, any system that does not induce any additional information is optimal for the designer. In it, there is no reason for the expert to engage in reputation-building, so his equilibrium payoff is u, the highest possible indeed.

The more interesting case is when $\rho > \rho^*$ (or $\lambda < 1$). The expert would like to spend as much time as possible in hiring ratings and the lowest possible in non-hiring ratings. Unless the rating system triggers incentives for the expert to manipulate reputation by providing t_L for problem H, it would be optimal for him, from a myopic point of view, to tell the truth whenever hired. Proposition 3 below derives an upper bound on the expert's probability of being hired and proves that such bound can be achieved.[21]

Proposition 3. *Given any prior belief $\rho \in (\rho^*, \bar{\rho})$, in any rating system inducing an informative equilibrium, there exists an upper bound on the expert's probability of being hired:*

$$\sum_{i \in \mathcal{I}} f_i^S \leq \frac{2\lambda - 1}{\lambda}.$$

This upper bound is achieved if (i) induced beliefs in all non-hiring ratings equal $\bar{\rho}$, (ii) induced beliefs in all hiring ratings equal ρ^* and (iii) the expert tells the truth whenever hired.

Proof. By Proposition 1, all non-hiring ratings have $\frac{f_i^B}{f_i^S} \leq 2\lambda$; so if $\mathcal{I}$ is the set of hiring ratings, $\frac{1-\sum_{i\in\mathcal{I}} f_i^B}{1-\sum_{i\in\mathcal{I}} f_i^S} \leq 2\lambda$, which rearranges to imply

$$2\lambda \sum_{i \in \mathcal{I}} f_i^S \leq 2\lambda - 1 + \sum_{i \in \mathcal{I}} f_i^B. \tag{11}$$

But equation (8) implies that consumers will only hire the expert in ratings where their posterior on a bad type is below ρ^*, so all hiring ratings i have $\frac{f_i^B}{f_i^S} \leq \lambda$, implying $\sum_{i\in\mathcal{I}} f_i^B \leq \lambda \sum_{i\in\mathcal{I}} f_i^S$. Substituting into the right-hand side above, this implies that a strategic type gets hired with probability $\sum_{i\in\mathcal{I}} f_i^S \leq \frac{2\lambda-1}{\lambda}$. This upper bound is achieved if (i) equations (8) and (11) bind; (ii) expert tells the truth in every hiring rating whenever hired. Note that these equations bind if every hiring rating i has $\rho_i = \rho^*$ and every non-hiring rating j has $\rho_j = \bar{\rho}$. □

We depict the expert's upper bound on the hiring probability in Fig. 4 below, for every customers' prior belief about expert being bad. The solid, black line represents the hiring probability under the one-shot interaction and the loosely dotted, blue line the bound from Proposition 3.

Proposition 3 implies that it suffices to have a binary system to achieve the maximum bound, if a binary system gives the right incentives for the expert to provide the correct treatment when hired. Without loss of generality, assume that he has no reason to provide inadequate treatment for problem L. The next proposition shows precisely that a binary system does not tempt the expert to manipulate reputation. The proof is in the Appendix, but the intuition is simple. With only two ratings, providing t_L upon the observation of problem H leads to a negative expert's payoff whenever hired without improving the expert's reputation, since with two ratings, it is already at its maximum.

Proposition 4. *Consider a binary rating system inducing an informative equilibrium. Playing truthfully in the hiring rating is always optimal for the expert.*

[21] Proposition 3 is an adaptation of Theorem 1 in Lorecchio and Monte (2023).

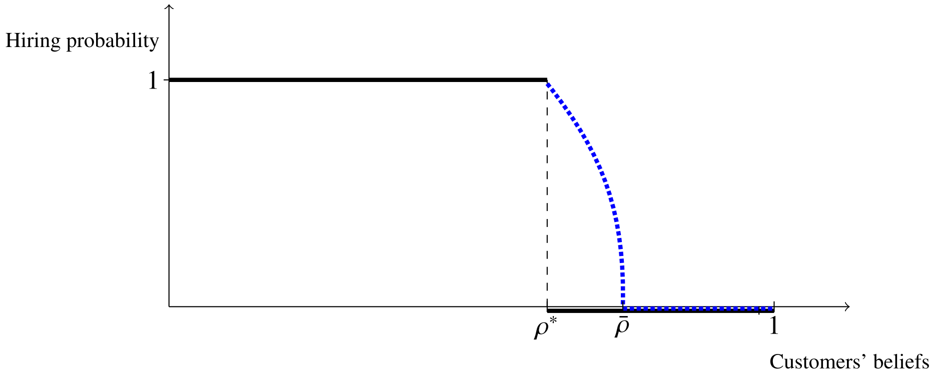

Fig. 4. Expert's hiring probabilities as functions of customers' (prior) beliefs. The solid black line gives the probability in the repetition of the one-shot case, or under full-censoring. The dotted, blue line corresponds to the upper bound on the hiring probability under any rating system inducing an informative equilibrium.

Analogous to the previous section, for every prior belief between ρ^* and $\bar{\rho}$, and any discount factor between zero and one, we can construct a sequence of ratings systems inducing an informative equilibrium in which the expert's payoff approaches the maximum value u. Unlike the construction of the systems in Theorem 1 however, it suffices here to consider binary rating systems. Along this sequence of systems, customers hire only in rating 2 and the expert tells the truth after being hired. Moreover, the construction induces posterior beliefs ρ^* and $\bar{\rho}$ in the non-hiring and the hiring ratings, respectively. Therefore, it reaches the upper bound derived in Proposition 3.

This construction is given in Theorem 2 below, the proof of which is in the Appendix. In constructing the systems, we set the initial rating to be the hiring and we let the transition probability out of the hiring rating, conditional on the observation of t_H, to be arbitrarily close to zero, while adjusting the transition out of the non-hiring rating so that the posterior beliefs are well-defined and equal to the desired values.

Theorem 2. *For any prior belief $\rho \in (\rho^*, \bar{\rho})$, any discount factor $\delta \in (0, 1)$, and $\forall \varepsilon > 0$ there is a sequence of binary rating systems $\mathcal{M}_t := (\{1, 2\}, \varphi_t, \varphi_0)$ and a number $T > 1$ such that (i) each element of the sequence induces an informative equilibrium; (ii) the expert's payoff approaches its first best: $V_t > u - \varepsilon, \forall t > T$. At each system, the expert tells the truth whenever hired and his probability of being hired equals the upper bound derived in Proposition 3.*

We note that this result is only possible if the discount factor is not exactly one. Intuitively, for $\delta < 1$, there is a difference between the customers and the expert with respect to the evaluation of the flow of payoffs throughout their interaction. While customers' payoff is a weighted combination of payoffs they get infinitely often in each rating, with weights corresponding to the long-run frequency of each rating, the expert' payoff is an average discounted combination of payoffs he gets in each rating. Therefore, for small chances of moving out of the hiring rating, and starting in the hiring rating, the expert's average discounted payoff is close to the hiring payoff u.

For the case in which the expert evaluates payoffs according to the stationary distributions, his payoff is bounded above since stationary distributions are bounded precisely in the same way as shown in Proposition 3. In this case, there is an even simpler construction: it is exactly the one described in Example 1. Such rating system achieves $\rho_1 = \bar{\rho}$ and $\rho_2 = \rho^*$ in equilibrium and, as a direct consequence of Bayes-plausibility to this environment, this yields the optimal ex-ante expected payoff. It implies that the expert is being hired and telling the truth infinitely often.

Combining this result with the results from Theorem 2, we see that the expert's equilibrium payoff is bounded away from zero, even for a sequence of discount factors approaching zero. Moreover, the bad reputation effect is eliminated, that is, the expert plays a truth-telling strategy whenever hired, and customers hire in at least one rating, even if $\rho > \rho^*$.

It is worth discussing the difference between the construction of the expert's and customers' optimal systems. Binary systems are sufficient here because (i) in non-hiring ratings, no additional information about the expert is generated and (ii) in hiring ratings there is no need to induce further learning.

5. Reducible equilibria

In this section, we discuss how our results extend without assuming that the bad expert visits all ratings infinitely often. Most importantly, we prove that any reducible equilibria obeys the same bound on the maximum induced belief as the bound derived in Proposition 1. But before doing so, we need some additional definitions.

Recall that any rating system $M = (M, \varphi, \varphi_0)$ and strategy profile $(\alpha, \beta^H, \beta^L)$ induce conditional transition matrices T^B and T^S. We can always partition the rating set into $n \geq 1$ irreducible (or ergodic) sets of ratings - denoted by $E_1, ..., E_n$ - and one transient, all from the bad expert's perspective. Every rating in the class $\cup_{m\in\{1,...,n\}} E_m$ of irreducible sets for the bad type of expert is a *recurrent* (or non-transient) rating for such type.[22] If the initial distribution φ_0 is such that there is a positive probability of the system eventually entering any irreducible set E_m, then all ratings $i \in E_m$ are *accessible* to the bad type, that is $f_i^B > 0$.[23] It turns out that, in equilibrium, all accessible ratings to the bad type are also accessible to the strategic type (Lemma 2 in the Appendix). In particular, if the entire rating set is irreducible for the bad type, then they are all accessible to all types and the steady-state distributions are independent of the initial distribution.[24] This was the case in previous sections. Thus, in this section we study equilibria in which some ratings are accessible to the strategic expert, but not to the bad type.

To provide a counterpart of the arguments in previous section, we highlight that the expected value of induced posteriors in any reducible equilibria must equal the prior. This is similar to the martingale property of posterior beliefs (Aumann et al., 1995) or Bayes-plausibility requirement in Bayesian persuasion models (Kamenica and Gentzkow, 2011), but applied to learning dynamics under automata.

Lemma 1. *Consider any rating system inducing an informative equilibrium. Let $\mathcal{A}^B$ represent the class of accessible sets to the bad type of expert. Define $\mathcal{A}^S$ as the class of accessible sets to the strategic expert. Let $f_i := \rho f_i^B + (1-\rho) f_i^S$, for each $i \in \mathcal{A}^B \cup \mathcal{A}^S$. Then*

$$\sum_{i \in \mathcal{A}^B \cup \mathcal{A}^S} f_i \rho_i = \sum_{i \in \mathcal{A}^B} f_i \rho_i = \rho.$$

The second equality follows because any rating that is accessible to the strategic type but not accessible to the bad type has its induced posterior belief equal to zero. From this lemma, there are two cases to consider: (i) $\rho \geq \bar{\rho}$ and (ii) $\rho < \bar{\rho}$.

In the first case, we claim that there can be no informative equilibria. Indeed, assume by way of contradiction that there exists at least one hiring rating in equilibrium. This means that there is at least one rating with an induced belief strictly lower than the prior (all hiring ratings must have belief no higher than ρ^*, and $\rho^* < \bar{\rho}$). Then Bayes plausibility would imply at least one induced belief strictly higher than ρ. The rating associated with such belief must be a non-hiring rating. We claim that, in equilibria with $\rho > \bar{\rho}$, all beliefs in non-hiring ratings must be weakly lower than ρ. The argument is as follows.

Suppose there is a rating i with $\rho_i > \rho$. This belief must have been formed through the weighted posterior of all possible histories in transient ratings that may have led to that rating. We need at least one of these histories to have had induced a posterior that is greater than the prior. This can not happen. For transient hiring ratings, due to incentive compatibility, (i) the belief is bounded above by ρ^*; (ii) the probability of the strategic expert transitioning into rating i is bounded below by the probability of the bad expert transitioning as well. The combination of a low belief in the transient hiring rating and the bound on the probability of the strategic expert transitioning to i leads to the posterior belief at theses transient ratings being at most $\bar{\rho}$. For transient non-hiring ratings, the posterior after leaving the rating must be equal to the prior entering that rating, since nothing is learned in non-hiring ratings. Therefore, a belief higher than $\rho \geq \bar{\rho}$ is never generated.

Having discussed the case in which $\rho \geq \bar{\rho}$, we now consider the case with $\rho < \bar{\rho}$. The next proposition extends the result about a belief upper bound in Proposition 1 to account for informative and reducible equilibria. It says that, for every accessible rating to the bad expert in an informative equilibrium, the highest induced posterior belief is no higher than the exactly same bound derived under irreducible equilibria. The proof is in the Appendix.

Proposition 5. *In any rating system inducing an informative equilibrium, for all accessible ratings i to the bad type,*

$$\rho_i \leq \bar{\rho}.$$

6. Conclusion

People rely on online reputation systems to hire experts, but such systems can generate wrong incentives for experts and mistrust from customers. In this paper, we consider the work by Ely and Välimäki (2003) regarding the stylized interaction between those players, but with an information intermediary. Without any informational restriction, the expert cannot avoid being rarely hired in equilibrium since either the belief about him being bad is eventually high enough that customers no longer hire, or the temptation to reveal himself by lying to a customer is so strong that market collapses.

With rating systems, we show that there is a bound on the equilibrium payoffs - which we characterize in terms of the posterior belief spread. The upper bound on induced beliefs does not depend on all ratings being issued infinitely often (that

[22] A rating j is a *consequent* of i for type $\omega \in \{B, S\}$ if $\tau_{ij}^{\omega,(n)} > 0$ for some $n \in \mathbb{N}$, where $\tau_{ij}^{\omega,(n)}$ is the (i, j)-entry of the n-th power of T^ω. In words, if there is a positive probability of going from i to j in a finite number of periods. A rating i is *transient* if it has at least one consequent j for which $\tau_{j,i}^{\omega,(n)} = 0$ for all $n \in \mathbb{N}$. Stating it differently, if there is a positive probability of leaving it and never returning. A rating i is *recurrent* if for every j that is a consequent of i, i is also a consequent of j. Note that every rating is either transient or recurrent. A well-known result states that, for any finite rating set and associated transition matrix, at least one irreducible subset always exists (Stokey and Lucas, 1989, Theorem 11.1).

[23] Note that all accessible ratings to type ω are recurrent, but the converse is not true.

[24] See Stokey and Lucas, 1989, Theorem 11.2.

is, if the system is reducible or irreducible). Nevertheless, bad reputation is avoidable: the expert tells the truth whenever hired - and gets hired even in the long run.

From the customers' perspective, we construct a finite and irreducible system to arbitrarily approach the equilibrium payoff upper bound (Theorem 1). In the optimal system, the interaction takes place most often at the extreme ratings, even though intermediary ones are important to further separate types. To curb the incentives for the expert to engage in reputation building at the cost of harming customers, the optimal system has to take into account the expert's discount factor. Specifically, the more he cares about frequently interacting with customers, the less sensible to the choice of treatment the downgrades or upgrades of his ratings should be.

From the expert's perspective, a binary and irreducible rating system is sufficient for approaching his optimal payoff (Theorem 2). This follows from non-hiring ratings being uninformative - no treatment is observed - and the expert only caring about being hired in equilibrium.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

We are thankful for conversations with Braz Camargo, Eduardo Faingold, Marcos Nakaguma, R. Vijay Krishna and Thomas Wiseman. We also thank the seminar participants at SBE 2021, Dehli Winter School 2021, NASMES 2022, Fourth BEAT Trobada, 12th Annual Conference of the Israeli Game Theory Society, Universitat de Barcelona and Florida State University. Caio Lorecchio acknowledges the financial support from the Spanish Ministry of Science and Innovation within the call *"Proyectos de Generación de Conocimiento"* (PID2021-128737 OB-I00), from Universitat de Barcelona within the call *"Ajuts per a Projectes d'Àrees Emergents"* (UB-AE-2022) and from the Agency for Management of Universities and Research Grants of the Government of Catalonia, within the call *"Suport a l'Activitat Científica dels Grups de Recerca"* (SGR2021-00862). Daniel Monte thanks the financial support from FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), Bolsa de Pesquia - Exterior 2021/06777-3 and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Appendix A. Proofs

Proof of Proposition 1. Let $\mathcal{O}_1$ be the set of non-hiring ratings i such that $\frac{f_i^B}{f_i^S} > 2\lambda$ and $\mathcal{O}_2$ the set of non-hiring ratings such that $\frac{f_i^B}{f_i^S} \leq 2\lambda$. Assume by way of contradiction that $\mathcal{O}_1$ is non-empty. With a slight abuse of notation, let $\tau_{i,C}^{\omega}$ be the transition chance from rating i to a set of ratings C conditional on $\omega \in \{B, S\}$. For every non-hiring rating, we eliminate the superscript ω because the transition chance will be uninformative.

We will use a well-known result for Markov processes about transitioning in and out of a partition of the space. If a state space is partitioned into two subsets, then in the steady state the probability of transitioning out of one subset to the other must equal the probability of transitioning into the former from the latter. Applying it to the environment considered in this paper, the steady-state probabilities must obey the following:

$$\frac{\sum_{i\in\mathcal{O}_1} f_i^B \left(\tau_{i,\mathcal{O}_2} + \tau_{i,\mathcal{I}}\right)}{\sum_{i\in\mathcal{O}_1} f_i^S \left(\tau_{i,\mathcal{O}_2} + \tau_{i,\mathcal{I}}\right)} = \frac{\sum_{i\in\mathcal{O}_2} f_i^B \tau_{i,\mathcal{O}_1} + \sum_{i\in\mathcal{I}} f_i^B \tau_{i,\mathcal{O}_1}^B}{\sum_{i\in\mathcal{O}_2} f_i^S \tau_{i,\mathcal{O}_1} + \sum_{i\in\mathcal{I}} f_i^S \tau_{i,\mathcal{O}_1}^S}. \tag{A.1}$$

The left-hand side exceeds 2λ by assumption since it is a weighted average of the ratios $\frac{f_i^B}{f_i^S}$, $i \in \mathcal{O}_1$ - each weight is $\frac{f_i^S(\tau_{1,\mathcal{O}_2}+\tau_{i,\mathcal{I}})}{\sum_{i\in\mathcal{O}_1} f_i^S(\tau_{1,\mathcal{O}_2}+\tau_{i,\mathcal{I}})}$. So it suffices to show that also $f_i^B \tau_{i,\mathcal{O}_1}^B \leq 2\lambda f_i^S \tau_{i,\mathcal{O}_1}^S$ $\forall i \in \mathcal{I}$. But $\tau_{i,\mathcal{O}_1}^S \geq \gamma_i \varphi_{i,\mathcal{O}_1}^H = \gamma_i \tau_{i,\mathcal{O}_1}^B$, so $\tau_{i,\mathcal{O}_1}^B \leq \frac{1}{\gamma_i} \tau_{i,\mathcal{O}_1}^S$. The ratio $\frac{1}{\gamma_i}$ is well-defined because there is a positive probability of t_H being played by the strategic expert in any hiring rating. Indeed, inspecting customers' expected payoff in any possible hiring rating i (equation (4)), we see that the expert must be willing to tell the truth in it with minimum probability, given by

$$\beta_i \geq \frac{1}{(u+w)} \left[w + \frac{\rho_i}{1-\rho_i} \left(\frac{w-u}{2} \right) \right] \Rightarrow \beta_i \geq \frac{w}{u+w}.$$

Because $\beta_i = \frac{1}{2}\beta_i^H + \frac{1}{2}\beta_i^L$, there will be a lower bound on the probability of choosing t_H when problem is H:

$$\frac{1}{2}\beta_i^H \geq \frac{w}{u+w} - \frac{1}{2}\beta_i^L \Rightarrow \beta_i^H \geq \frac{w-u}{w+u}.$$

This lower bound then leads to γ_i being positive:

$$\begin{aligned}\gamma_i &= \frac{1}{2}\beta_i^H + \frac{1}{2}(1-\beta_i^L),\\ &\geq \frac{1}{2}\left[\frac{w-u}{w+u}\right],\\ &> 0.\end{aligned}$$

Therefore, substituting $\tau_{i,\mathcal{O}_1}^B \leq \frac{1}{\gamma_i}\tau_{i,\mathcal{O}_1}^S$ into equation (8), we have

$$\begin{aligned} i \in \mathcal{I} \ \Rightarrow \ f_i^B \tau_{i,\mathcal{O}_1}^B &\leq \lambda \left[\frac{\beta_i - (1-\beta_i)\frac{w}{u}}{\gamma_i}\right] f_i^S \tau_{i,\mathcal{O}_1}^S,\\ &\leq 2\lambda \left[\frac{2\beta_i - 1}{2\gamma_i}\right] f_i^S \tau_{i,\mathcal{O}_1}^S,\\ &= 2\lambda \left[\frac{\beta_i^H - (1-\beta_i^L)}{\beta_i^H + (1-\beta_i^L)}\right] f_i^S \tau_{i,\mathcal{O}_1}^S.\end{aligned}$$

The second inequality follows from $\frac{w}{u} \geq 1$; the equality follows from the definition of β_i and γ_i. Therefore, $f_i^B \tau_{i,\mathcal{O}_1}^B \leq 2\lambda f_i^S \tau_{i,\mathcal{O}_1}^S$ $\forall i \in \mathcal{I}$, as desired. □

Proof of Theorem 1. We proceed through the following steps: (I) We first choose M_0 high enough so that the given prior belief is above a threshold $\underline{\rho}(M_0) \leq \rho^*$ (defined below). Given M_0, we construct a system with M_0 ratings and a transition rule such that, as the probability of leaving extreme ratings approaches zero, we get $\rho_1 \approx \bar{\rho}$, $\rho_2 \approx \rho^*$ and $\rho_{M_0} \approx \underline{\rho}(M_0)$. In this system, hiring is optimal for customers in every rating except rating 1, if the strategic expert optimally provides the adequate treatment whenever hired. (II) Given that customers' strategy is incentive compatible at the induced beliefs, we adjust parameters so that truth-telling is indeed optimal whenever hired, for every $\delta < 1$. (III) with the adjusted parameters, we show that the induced payoff is ε-close to customers' maximum equilibrium payoff.

I - Choosing the number of ratings and the transition rule. Consider some $M_0 \geq 2 \in \mathbb{N}$ big enough so that $\rho > \underline{\rho}(M_0)$, where $\underline{\rho}(M_0)$ is given by

$$\underline{\rho}(M_0) := \frac{\rho^*\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\rho^*+\rho^*\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}. \tag{A.2}$$

As long as $\rho \in (\underline{\rho}(M_0), \bar{\rho})$, for a rating system with M_0 ratings, an equilibrium with $\rho_1 \leq \bar{\rho}$ and $\rho_{M_0} \geq \underline{\rho}(M_0)$ is at least possible. From Proposition 2, it suffices to have only rating 1 as a non-hiring rating in this system. For now, assume that the strategic expert optimally tells the truth whenever hired (we will deal with the incentive compatibility of this player later on).

We will use the following quantities to design the appropriate transition rule:

$$\begin{aligned} \varphi_+^H &:= \left(\frac{1-\rho^*}{2}\right)\psi \qquad & \varphi_-^L &:= 0,\\ \varphi_+^L &:= \left(\frac{1+\rho^*}{2}\right)\psi \qquad & \varphi_-^H &:= \psi,\end{aligned} \tag{A.3}$$

where $\psi \in (0,1)$ will be chosen appropriately throughout the proof.

Consider the following transitions. At rating 1, there is a random exit probability τ to rating 2. At any $m \notin \{1, M_0\}$, (a) whenever treatment t_H is observed, the rating is upgraded with probability φ_+^H and downgraded with probability φ_-^H (with ψ small enough to have $\varphi_+^H + \varphi_-^H \leq 1$); (b) whenever t_L is observed, the rating is upgraded with probability φ_+^L and does not change with probability $1-\varphi_+^L$. At rating M_0, treatment t_H leads to a downgrade with probability κ and treatment t_L does not change the rating. Similar to ψ, the parameters τ and κ are chosen appropriately through the proof.

Observe that, under a truth-telling strategy, the transition probabilities generate the following properties for every intermediary rating $m \notin \{1, M_0\}$: (a) $\frac{\tau_{mm+1}^B}{\tau_{m+1m}^B} = \frac{1-\rho^*}{2}$; (b) $\tau_{mm+1}^S = \tau_{m+1m}^S$ and (c) $\varphi_+^L > \varphi_+^H$. In words, property (a) implies that the bad expert gets downgraded with more probability than upgraded and the ratio of such probabilities equals the minimum chance of the strategic expert playing t_H in hiring ratings. Property (b) implies that the strategic expert has same chances of being upgraded or downgraded - $\frac{\psi}{2}$ - under a truth-telling strategy. If he plays the correct treatment for a major problem with the minimum probability that justifies hiring, he is upgraded with the same probability of a bad expert. Property (c) means that treatment t_L leads to an upgrade with higher probability than t_H.

With this transition rule, we have the following identities.

$$f_1^B \tau = \psi f_2^B, \qquad f_{m+1}^B = \left(\frac{1-\rho^*}{2}\right) f_m^B, \qquad f_{M_0}^B \kappa = \left(\frac{1-\rho^*}{2}\right) \psi f_{M_0-1}^B;$$
$$f_1^S \tau = \left(\frac{\psi}{2}\right) f_2^S, \qquad f_{m+1}^S = f_m^S, \qquad f_{M_0}^S \left(\frac{\kappa}{2}\right) = \left(\frac{\psi}{2}\right) f_{M_0-1}^S.$$

We can find the values of f_1^B and f_1^S through the following system of equations.

$$1 = f_1^B + \frac{\tau}{\psi} f_1^B + \left(\frac{1-\rho^*}{2}\right)\left[\frac{\tau}{\psi}\right] f_1^B + \dots + \left(\frac{1-\rho^*}{2}\right)^{M_0-3}\left[\frac{\tau}{\psi}\right] f_1^B + \frac{\tau}{\kappa}\left(\frac{1-\rho^*}{2}\right)^{M_0-2} f_1^B,$$
$$1 = f_1^S + 2\left[\frac{\tau}{\psi}\right] f_1^S + 2\left[\frac{\tau}{\psi}\right] f_1^S + \dots + 2\left[\frac{\tau}{\psi}\right] f_1^S + 2\left(\frac{\tau}{\kappa}\right) f_1^S.$$

The solutions are

$$f_1^B = \frac{1}{1 + \frac{\tau}{\psi}\left(\frac{1-\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\left(\frac{1-\rho^*}{2}\right)}\right) + \frac{\tau}{\kappa}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}, \tag{A.4}$$

$$f_1^S = \frac{1}{1 + 2\left(\frac{\tau}{\psi}\right)[M_0-2] + 2\left(\frac{\tau}{\kappa}\right)}. \tag{A.5}$$

We want to set $\frac{\tau}{\psi} \approx 0$ but keep $\frac{\tau}{\kappa} > 0$. This will lead to intermediary ratings been rarely visited by both types in equilibrium. Indeed, because f_2^B and f_2^S are

$$f_2^B = \frac{\frac{\tau}{\psi}}{1 + \frac{\tau}{\psi}\left(\frac{1-\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\left(\frac{1-\rho^*}{2}\right)}\right) + \frac{\tau}{\kappa}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}, \tag{A.6}$$

$$f_2^S = \frac{2\left(\frac{\tau}{\psi}\right)}{1 + 2\left(\frac{\tau}{\psi}\right)[M_0-2] + 2\left(\frac{\tau}{\kappa}\right)}. \tag{A.7}$$

As $\frac{\tau}{\psi} \approx 0$, we will have $f_2^B \approx f_2^S \approx 0$. From our derivation of the stationary probabilities, it will also be the case $f_m^B \approx f_m^S \approx 0$ for $m \in \{3, \dots, M_0-1\}$. However, for rating M_0,

$$f_{M_0}^B = \frac{\frac{\tau}{\kappa}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1 + \frac{\tau}{\psi}\left(\frac{1-\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\left(\frac{1-\rho^*}{2}\right)}\right) + \frac{\tau}{\kappa}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}, \tag{A.8}$$

$$f_{M_0}^S = \frac{2\left(\frac{\tau}{\kappa}\right)}{1 + 2\left(\frac{\tau}{\psi}\right)[M_0-2] + 2\left(\frac{\tau}{\kappa}\right)}. \tag{A.9}$$

So even if $\frac{\tau}{\psi} \approx 0$, as long as $\frac{\tau}{\kappa} > 0$, those probabilities will be positive. We also want to set $\frac{f_1^B}{f_1^S} = 2\lambda$. Using equations (A.4) and (A.5), we can find the value of the ratio $\frac{\tau}{\kappa}$ that leads to ratio 2λ for $\frac{\tau}{\psi} \approx 0$. This value is

$$\frac{\tau}{\kappa} = \frac{\lambda - \frac{1}{2}}{1 - \lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}. \tag{A.10}$$

Because $\rho < \bar{\rho}$ implies $\lambda > 1/2$, the numerator is positive. Because $\rho > \underline{\rho}(M_0)$, the denominator will be positive as well. Thus, this ratio is well defined. With this system, we also achieve $\rho_{M_0} = \underline{\rho}(M_0)$. Indeed, note from equations (A.8) and (A.9) that the relative frequency $\frac{f^B_{M_0}}{f^S_{M_0}}$ is

$$\begin{aligned}\frac{f^B_{M_0}}{f^S_{M_0}} &= \frac{1}{2}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}\left[\frac{1+2\left(\frac{\tau}{\kappa}\right)}{1+\frac{\tau}{\kappa}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right],\\ &= \frac{1}{2}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}\left[\frac{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}+2\lambda-1}{1-\frac{1}{2}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right],\\ &= \frac{1}{2}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}2\lambda\left[\frac{1-\frac{1}{2}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\frac{1}{2}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right],\\ &= \left(\frac{1-\rho^*}{2}\right)^{M_0-2}\lambda.\end{aligned}$$

Recalling that $\rho_{M_0} = \frac{\rho f^B_{M_0}}{(1-\rho) f^S_{M_0}}$ and $\lambda = \frac{(1-\rho)\rho^*}{\rho(1-\rho^*)}$, simple manipulation shows that

$$\rho_{M_0} = \frac{\rho^*\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\rho^*+\rho^*\left(\frac{1-\rho^*}{2}\right)^{M_0-2}} := \underline{\rho}(M_0).$$

II - adjusting parameters so that truth-telling is optimal for every $\delta < 1$. We have assumed as an equilibrium candidate a strategy profile in which customers hire in every rating except 1 and expert plays the truth-telling strategy whenever hired. It remains to adjust parameters so that the profile is an equilibrium. Note that customers' strategy is incentive compatible. At rating 1 not hiring is optimal indeed, because $\frac{f^B_1}{f^S_1} = 2\lambda > \lambda$. In rating 2, if expert provides the right treatment, then hiring is optimal, because $\frac{f^B_2}{f^S_2} = \lambda$, as it can be seen from equations (A.6) and (A.7). For every other rating $m \in \{3, ..., M_0\}$, we will have

$$\frac{f^B_m}{f^S_m} = \left(\frac{1-\rho^*}{2}\right)^{m-2}\lambda < \lambda,$$

so hiring is optimal in all of them.

Before we derive parameter values for which truth-telling is optimal for the strategic expert, we claim that expert's continuation value in every rating is bounded below by $-w$. To see this, let $\underline{m} \in \arg\min\{V_m : 1 \le m \le M_0\}$. From equations (1) and (5), the continuation value at $\underline{m}$ is such that

$$\begin{aligned}V_{\underline{m}} &= (1-\delta)\left\{\alpha_{\underline{m}}\left[\beta_{\underline{m}}u-(1-\beta_{\underline{m}})w\right]\right\}+\delta\sum_{m=1}^{M_0}\tau^S_{\underline{m},m'}V_{m'},\\ &\ge (1-\delta)(-w)+\delta V_{\underline{m}}.\end{aligned}$$

This implies that $V_{\underline{m}} \ge -w$; so expert's continuation value in every rating is bounded below by $-w$. With that in mind, we start the analysis with M_0. If the expert sticks to the prescribed equilibrium strategy, we have

$$V_{M_0} = (1-\delta)u+\delta\left[\left(1-\frac{\kappa}{2}\right)V_{M_0}+\frac{\kappa}{2}V_{M_0-1}\right].$$

Subtracting V_{M_0-1} on both sides and rearranging,

$$\begin{aligned}V_{M_0}-V_{M_0-1} &= \left(\frac{1-\delta}{1-\delta+\frac{\delta\kappa}{2}}\right)[u-V_{M_0-1}],\\ &\le \left(\frac{1-\delta}{1-\delta+\frac{\delta\kappa}{2}}\right)[u+w]. \qquad (A.11)\end{aligned}$$

If expert provides the correct treatment t_H at problem H, he gets

$$(1-\delta)u + \delta\left\{\kappa\left[V_{M_0-1} - V_{M_0}\right] + V_{M_0}\right\},$$

and if he provides treatment t_L instead, he gets

$$(1-\delta)(-w) + \delta V_{M_0}.$$

Thus, truth-telling is optimal if $\frac{1-\delta}{\delta}(u+w) > \kappa[V_{M_0} - V_{M_0-1}]$. But from equation (A.11),

$$\kappa[V_{M_0} - V_{M_0-1}] \le \kappa\left(\frac{1-\delta}{1-\delta+\frac{\delta\kappa}{2}}\right)[u+w].$$

Thus, as long as $\kappa < 2\left(\frac{1-\delta}{\delta}\right)$, telling the truth at rating M_0 will be optimal indeed.

Let us check now under what conditions telling the truth is optimal at rating M_0-1. We have

$$V_{M_0-1} = (1-\delta)u + \delta\left\{\frac{\psi}{2}\left[V_{M_0-2} - V_{M_0-1}\right] + V_{M_0-1} + \frac{\psi}{2}\left[V_{M_0} - V_{M_0-1}\right]\right\}.$$

Subtracting V_{M_0-2} on both sides and rearranging,

$$\begin{aligned} V_{M_0-1} - V_{M_0-2} &= \left(\frac{1-\delta}{1-\delta+\delta\frac{\psi}{2}}\right)\left[u - V_{M_0-2}\right] + \left(\frac{\delta\frac{\psi}{2}}{1-\delta+\delta\frac{\psi}{2}}\right)[V_{M_0} - V_{M_0-1}], \\ &\le (u+w), \end{aligned}$$

because $\psi > 0$, $V_{M_0-2} \ge -w$ as well as $V_{M_0} - V_{M_0-1} \le u+w$. If the expert provides the correct treatment t_H at problem H, he gets

$$(1-\delta)u + \delta\left\{\varphi_-^H\left[V_{M_0-2} - V_{M_0-1}\right] + V_{M_0-1} + \varphi_+^H\left[V_{M_0} - V_{M_0-1}\right]\right\}.$$

And if he provides treatment t_L instead, he gets

$$(1-\delta)(-w) + \delta\left\{V_{M_0-1} + \varphi_+^L\left[V_{M_0} - V_{M_0-1}\right]\right\}.$$

Therefore, truth-telling is optimal if

$$\frac{1-\delta}{\delta}(u+w) > (\varphi_+^L - \varphi_+^H)\Big[V_{M_0} - V_{M_0-1}\Big] + \varphi_-^H\Big[V_{M_0-1} - V_{M_0-2}\Big]. \tag{A.12}$$

From equation (A.11), the above discussion, the fact that $\varphi_+^L - \varphi_+^H = \rho^*\psi > 0$ and $\varphi_-^H = \psi$, the right hand side of this equation must be bounded above by

$$\psi\left[\frac{(1-\delta)(1+\rho^*) + \delta\frac{\kappa}{2}}{1-\delta+\delta\frac{\kappa}{2}}\right](u+w).$$

From the already derived bound for κ, we get that as long as

$$\psi < \frac{2}{2+\rho^*}\left(\frac{1-\delta}{\delta}\right), \tag{A.13}$$

equation (A.12) is satisfied. Assume this is the case. We now proceed by induction. Suppose we have proved that truth-telling is optimal for $M_0, M_0-1, \ldots, M_0-m$, with $m \le M_0-3$. We would like to show that truth-telling is optimal for $M_0-(m-1)$. We have

$$V_{M_0-m+1} = (1-\delta)u + \delta\left\{\frac{\psi}{2}[V_{M_0-m+2} - V_{M_0-m+1}] + V_{M_0-m+1} + \frac{\psi}{2}[V_{M_0-m} - V_{M_0-m+1}]\right\}.$$

Subtracting V_{M_0-m+2} on both sides and rearranging,

$$V_{M_0-m+1} - V_{M_0-m+2} = \left(\frac{1-\delta}{1-\delta+\delta\frac{\psi}{2}}\right)\left[u - V_{M_0-m+2}\right] + \left(\frac{\delta\frac{\psi}{2}}{1-\delta+\delta\frac{\psi}{2}}\right)[V_{M_0-m} - V_{M_0-m}].$$

The right hand side of the above equation is at most $u+w$ because $\psi > 0$ as well as $V_{M_0-m+2} \ge -w$ and $V_{M_0-m} - V_{M_0-m+1} \le u+w$. From the same argument in previous paragraphs, truth-telling is optimal if

$$\frac{1-\delta}{\delta}(u+w) \geq (\varphi_{+}^{L}-\varphi_{+}^{H})[V_{M_0-m+1}-V_{M_0-m+2}]+\varphi_{-}^{H}[V_{M_0-m}-V_{M_0-m+1}].$$

The right hand side of the above equation is at most $\psi(1+\rho^*)[u+w]$, because $[V_{M_0-m}-V_{M_0-m+1}] \leq u+w$ and $[V_{M_0-m+1}-V_{M_0-m+2}] \leq u+w$. As long as $\psi \leq \frac{1}{1+\rho^*}\left(\frac{1-\delta}{\delta}\right)$, the above inequality is satisfied. This will be the case if ψ is lower than the bound from equation (A.13).

III - setting parameters so that the customers' equilibrium payoff is arbitrarily close to the maximum payoff. For every $t \in \mathbb{N}$, consider the following values:

$$\kappa_t = \frac{1}{t(t+1)^2}, \qquad \psi_t = \frac{2}{2+\rho^*}\left(\frac{1}{t+1}\right), \qquad \tau_t = \frac{1}{(t+1)^3}\left[\frac{\lambda-\frac{1}{2}}{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right].$$

Observe that

$$\frac{\tau_t}{\kappa_t} = \frac{t}{t+1}\left[\frac{\lambda-\frac{1}{2}}{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right], \qquad \frac{\tau_t}{\psi_t} = \frac{2+\rho^*}{2}\left(\frac{1}{t+1}\right)^2\left[\frac{\lambda-\frac{1}{2}}{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right].$$

This means that $\frac{\tau_t}{\psi_t} \to 0$ as $t \to \infty$, but $\frac{\tau_t}{\kappa_t}$ converges to equation (A.10). Given $\delta \in (0,1)$, and passing to a subsequence if necessary, we can define a sequence of parameter values such that, for each element of this sequence, $\kappa_t < 2\left(\frac{1-\delta}{\delta}\right)$, $\psi_t < \frac{2}{2+\rho^*}\left(\frac{1-\delta}{\delta}\right)$ and $\tau_t < 1$. Moreover, it will be the case that for every $t \in \mathbb{N}$,

$$\lambda < \frac{f_1^B}{f_1^S} \leq 2\lambda.$$

To see that this is true, note that, under the described parameter values,

$$\frac{f_1^B}{f_1^S} = \frac{1+2\left(\frac{\tau_t}{\psi_t}\right)[M_0-2]+2\left(\frac{\tau_t}{\kappa_t}\right)}{1+\frac{\tau_t}{\psi_t}\left(\frac{1-\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\left(\frac{1-\rho^*}{2}\right)}\right)+\frac{\tau_t}{\kappa_t}\left(\frac{1-\rho^*}{2}\right)^{M_0-2}} \leq 2\lambda,$$

$$\Leftrightarrow \frac{2+\rho^*}{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\left\{M_0-2-\lambda\left(\frac{1-\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\left(\frac{1-\rho^*}{2}\right)}\right)\right\} \leq 2(t+1).$$

Thus, whatever the value of M_0 is, there will be a t big enough to satisfy the above inequality. Likewise, $\frac{f_1^B}{f_1^S} > \lambda$ if and only if

$$\frac{\lambda-\frac{1}{2}}{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\left\{\frac{2+\rho^*}{(t+1)^2}\left[M_0-2-\frac{\lambda}{2}\left(\frac{1-\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\left(\frac{1-\rho^*}{2}\right)}\right)\right]+\right.$$
$$\left.+\frac{t}{t+1}\left[2-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}\right]\right\} > \lambda-1.$$

For t big enough, the right hand side of the above equation approaches

$$\left(\lambda-\frac{1}{2}\right)\left[\frac{2-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}{1-\lambda\left(\frac{1-\rho^*}{2}\right)^{M_0-2}}\right] > \lambda-\frac{1}{2} > \lambda-1.$$

With such sequence of parameter values and equilibria, we have

$$\lim_{t\to\infty} v = f_{M_0}\left[\rho_{M_0}\left(\frac{u-w}{2}\right)+(1-\rho_{M_0})u\right],$$
$$= \frac{\bar{\rho}-\rho}{\bar{\rho}-\underline{\rho}(M_0)}\left\{\underline{\rho}(M_0)\left(\frac{u-w}{2}\right)+[1-\underline{\rho}(M_0)]u\right\}.$$

The arbitrary approximation to the upper bound $u\left(\frac{\bar{\rho}-\rho}{\bar{\rho}}\right)$ comes by considering a large enough $M \geq M_0$ and the discussed sequences of parameters $\{\kappa_t, \psi_t, \tau_t\}_{t\in\mathbb{N}}$ given M. This shows the existence of a sequence of finite, irreducible, and incentive-compatible rating systems approaching the customers' maximum equilibrium payoff. $\square$

Proof of Proposition 4. A strategy with $\beta_2^H = 1$ is optimal if

$$(1-\delta)u + \delta[\varphi_{21}^H V_1 + \varphi_{22}^H V_2] \geq \beta_2^H\{(1-\delta)u + \delta[\varphi_{21}^H V_1 + \varphi_{22}^H V_2]\} + (1-\beta_2^H)\{(1-\delta)(-w) + \delta[\varphi_{21}^L V_1 + \varphi_{22}^L V_2]\},$$

which simplifies to:

$$(1-\delta)u + \delta[\varphi_{21}^H V_1 + \varphi_{22}^H V_2] \geq (1-\delta)(-w) + \delta[\varphi_{21}^L V_1 + \varphi_{22}^L V_2],$$

or $(1-\delta)(u+w) \geq \delta(\varphi_{21}^H - \varphi_{21}^L)[V_2 - V_1]$. To see this is indeed true, note that, from the continuation value in rating 2,

$$V_2 = (1-\delta)u + \delta\left[\tau_{21}^S V_1 + \tau_{22}^S V_2\right] \Rightarrow V_2 - V_1 = \frac{(1-\delta)[u - V_1]}{1-\delta+\delta\tau_{21}^S}.$$

Thus,

$$\begin{aligned}
\delta(\varphi_{21}^H - \varphi_{21}^L)[V_2 - V_1] &\leq \delta\varphi_{21}^H[V_2 - V_1],\\
&= (1-\delta)[u - V_1]\left\{\frac{\delta\varphi_{21}^H}{1-\delta+\delta\tau_{21}^S}\right\},\\
&< (1-\delta)[u + w]\left\{\frac{\delta\frac{1}{2}\varphi_{21}^H}{1-\delta+\delta\tau_{21}^S}\right\},\\
&< (1-\delta)(u+w).
\end{aligned}$$

The second inequality follows from $V_1 > 0 > \frac{u-w}{2}$ (the payoff in rating 1 is positive, provided that there is a positive chance of moving into the hiring rating 2); the third inequality follows from the expert's transition probability from 2 to 1 being $\tau_{21}^S = \frac{1}{2}\varphi_{21}^H + \frac{1}{2}\varphi_{21}^L \geq \frac{1}{2}\varphi_{21}^H$. $\square$

Proof of Theorem 2. The system starts in rating 2. There is a random exit probability τ from rating 1 and a random exit probability κ from rating 2 upon the observation of t_H. Whenever t_L is played in rating 2, the system stays in 2. If the strategy profile of not hiring in rating 1 and telling the truth in rating 2 is optimal, then the transition matrices will be

$$T^B = \begin{array}{c} \\ 1 \\ 2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 1 & \quad 2 \end{array} \\ \begin{pmatrix} 1-\tau & \tau \\ \kappa & 1-\kappa \end{pmatrix} \end{array} \qquad T^S = \begin{array}{c} \\ 1 \\ 2 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 1 & \quad 2 \end{array} \\ \begin{pmatrix} 1-\tau & \tau \\ \frac{\kappa}{2} & 1-\frac{\kappa}{2} \end{pmatrix} \end{array}$$

This leads to the following values for the invariant distributions

$$f_1^B = \frac{\kappa}{\kappa+\tau}, \qquad f_1^S = \frac{\kappa}{\kappa+2\tau}.$$

We need to find values of κ and τ such that $\frac{f_1^B}{f_1^S} = 2\lambda$. This is satisfied when

$$\frac{\tau}{\kappa} = \frac{\lambda - \frac{1}{2}}{1-\lambda}. \tag{A.14}$$

From Assumption 1, we know that $\lambda > \frac{1}{2}$; from the values of λ for which we are considering in this section, $\lambda < 1$.

We can derive the expert's continuation values in each rating, given any $\delta \in [0,1)$, by solving the following system of equations.

$$\begin{aligned}
V_1 &= \delta\left[(1-\tau)V_1 + \tau V_2\right],\\
V_2 &= (1-\delta)u + \delta\left[\frac{\kappa}{2}V_1 + \left(1-\frac{\kappa}{2}\right)V_2\right].
\end{aligned}$$

This leads to

$$V_1 = \left[\frac{\delta\tau}{1-\delta+\delta\left(\tau+\frac{\kappa}{2}\right)}\right]u \qquad V_2 = \left[\frac{1-\delta+\delta\tau}{1-\delta+\delta\left(\tau+\frac{\kappa}{2}\right)}\right]u.$$

Because the system starts in rating 2, the expert's equilibrium payoff is

$$V = \left[\frac{1-\delta+\delta\tau}{1-\delta+\delta\left(\tau+\frac{\kappa}{2}\right)} \right] u. \tag{A.15}$$

For every $t \in \mathbb{N}$, consider the following values:

$$\kappa_t = \frac{1}{t+1}, \qquad \tau_t = \left(\frac{\lambda - \frac{1}{2}}{1-\lambda} \right) \frac{1}{t+1}.$$

Passing to a subsequence if necessary, for every number t, the associated transition probabilities are well-defined values, leading to posterior beliefs ρ^* and $\bar{\rho}$ in the hiring and the non-hiring rating, respectively. Substituting for such values in equation (A.15) and rearranging, we obtain

$$V = \left[\frac{(1-\delta)\left(\frac{1-\lambda}{\lambda}\right) + \frac{\delta}{2(t+1)}\left(\frac{2\lambda-1}{\lambda}\right)}{(1-\delta)\left(\frac{1-\lambda}{\lambda}\right) + \frac{\delta}{2(t+1)}} \right] u.$$

The arbitrary approximation to u obtains by taking t to be big enough in the above equation. □

Lemma 2. *In any equilibrium, every accessible rating to the bad type of expert is also accessible to the strategic type.*

Proof. Assume by way of contradiction that there exists a rating i for which $f_i^B > 0$, but $f_i^S = 0$. As such, i must be a non-hiring rating. Moreover, this rating must belong to an irreducible set E for the bad type of expert. There are two cases to consider.

In the fist case, i is transient for the strategic type. This means that there exists at least one $j \neq i$ such that $\tau_{ij}^S > 0$, but $\tau_{ji}^{S,(n)} = 0$ for all $n \in \mathbb{N}$. Because E is an irreducible set for the bad expert, there exists $k \in E$ such that $\tau_{ki}^{B,(n)} > 0$ for some $n \in \mathbb{N}$. But then $\tau_{ki}^{B,(n)} \tau_{i,j}^B = \tau_{ki}^{B,(n)} \tau_{ij}^S > 0$. This means that j is a consequent of k for the bad type as well. We claim that it must also be the case that $\tau_{ji}^{B,(n)} = 0$ for all n. Indeed, observe that

$$\begin{aligned}
\tau_{ji}^S &= \alpha_j[\gamma_j \varphi_{ji}^H + (1-\gamma_j)\varphi_{ji}^L] + (1-\alpha_j)\varphi_{ji}^{Out}, \\
&= \gamma_j[\alpha_j \varphi_{ji}^H + (1-\alpha_j)\varphi_{ji}^{Out}] + (1-\gamma_j)[\alpha_j \varphi_{ji}^L + (1-\alpha_j)\varphi_{ji}^{Out}], \\
&= \gamma_j \tau_{ji}^B + (1-\gamma_j)[\alpha_j \varphi_{ji}^L + (1-\alpha_j)\varphi_{ji}^{Out}].
\end{aligned}$$

If j is a non-hiring rating, then $\tau_{ki}^S = \tau_{ki}^B = \varphi_{ki}^{Out}$. If j is a hiring rating, then $\tau_{ji}^S \geq \underline{\gamma}\varphi_{ji}^H = \underline{\gamma}\tau_{ji}^B$, where $\underline{\gamma} := \frac{1-\rho^*}{2} > 0$ is a lower bound on the probability of the strategic expert providing the high-cost treatment, in any hiring rating. Thus, $\tau_{ji}^B \leq \frac{1}{\underline{\gamma}}\tau_{ji}^S$. In matrix notation, $T^B \geq \text{diag}(\underline{\gamma}^{-1})T^S$, where $\text{diag}(\underline{\gamma}^{-1}) = \frac{1}{\underline{\gamma}} I_M$. Then

$$\begin{aligned}
(T^B)^n &\leq (\text{diag}(\underline{\gamma}^{-1})T^S)^n, \\
&= \text{diag}(\underline{\gamma}^{-n})(T^S)^n.
\end{aligned}$$

In particular, $\tau_{ji}^{B,(n)} \leq \underline{\gamma}^{-n}\tau_{ji}^{S,(n)}$. Therefore, $\tau_{ji}^{B,(n)} = 0$ for all $n \in \mathbb{N}$. But this means that k is transient for the bad type, which is a contradiction.

In the second case, i is recurrent for the strategic type. Then it must be that the initial distribution assigns zero probability to the irreducible set D for the strategic expert that i belongs to, and all transient ratings to which there is a positive probability of transitioning into D eventually. If $E \subseteq D$, a contradiction follows, so assume otherwise.

Take any $k \in E$ such that $k \notin D$. Because E is irreducible for the bad type, i is a consequent of k for the bad type, or $\tau_{ki}^{B,(n)} > 0$ for some $n \in \mathbb{N}$. But then $\tau_{k,i}^{S,(n)} \geq \underline{\gamma}\tau_{ki}^{B,(n)} > 0$, so i is also a consequent of k for the strategic type. Because the initial distribution either places positive probability on the equilibria starting in E or on transient ratings that transition to E eventually, it must be that $f_i^S > 0$, also a contradiction. □

Proof of Proposition 5. Let $\mathcal{A}^B$ be set of accessible ratings to the bad type of expert. Lemma 2 implies that $f_i^S > 0$ for all $i \in \mathcal{A}^B$. We can always partition $\mathcal{A}^B$ into hiring ratings $\mathcal{I}^B$ and non-hiring ratings $\mathcal{O}^B$.

Following the footsteps of the proof of Proposition 1, let $\mathcal{O}_1^B$ be the set of non-hiring ratings i such that $\frac{f_i^B}{f_i^S} > 2\lambda$ and $\mathcal{O}_2^B$ the set of non-hiring ratings such that $\frac{f_i^B}{f_i^S} \leq 2\lambda$. Note that, if $\rho < \bar{\rho}$, then there has to exist a rating in $\mathcal{O}_1^B$ that is a consequent of a rating in $\mathcal{I}^B$. Otherwise, any rating in $\mathcal{O}_1^B$ has a posterior no higher than the greatest posterior updated with the information that the system has entered $\mathcal{O}_1^B$. But from all transient ratings, either the information conveyed once the system enters $\mathcal{O}_1^B$ renders a posterior no lower than $\bar{\rho}$, or nothing new is learned (if the transient rating was a hiring rating). Thus, without being connected to the set of hiring ratings for the bad expert, all ratings in $\in \mathcal{O}_1^B$ would have $\rho_i \leq \bar{\rho}$ or $\frac{f_i^B}{f_i^S} \leq 2\lambda$ as well. Note in addition that there has to exist a rating in $\mathcal{I}^B$ that is a consequent of a rating in $\mathcal{O}_1^B$; otherwise, $\mathcal{O}_1^B$ would absorb all flow from $\mathcal{I}^B$.

For the strategic type, we also need to take into account non-hiring ratings for which $f_i^S > 0$ but $f_i^B = 0$; we will label this set as $\mathcal{O}^S$. Similarly, the set of hiring ratings for the strategic type is the partition between $\mathcal{I}^B$ and $\mathcal{I}^S$, where the latter is the set of hiring ratings that are not accessible to the bad type.

On the one hand, the steady-state probabilities for the strategic type must obey:

$$\sum_{i\in\mathcal{O}_1^B} f_i^S\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}+\tau_{i,\mathcal{O}^S}+\tau_{i,\mathcal{I}^S}\right)=\sum_{i\in\mathcal{O}_2^B} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} f_i^S\tau^S_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{O}^S} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^S} f_i^S\tau^S_{i,\mathcal{O}_1^B}.$$

The transition from every $i \in \mathcal{O}_1^B$ into $\mathcal{I}^S$ or $\mathcal{O}^S$ must be zero; otherwise, there would be a positive probability of the bad type of expert leaving recurrent rating i to non-accessible sets infinitely often, which is a contradiction. Thus, we obtain the following inequality:

$$\sum_{i\in\mathcal{O}_1^B} f_i^S\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}\right)\geq\sum_{i\in\mathcal{O}_2^B} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} f_i^S\tau^S_{i,\mathcal{O}_1^B}. \tag{A.16}$$

On the other hand, the steady-state probabilities for the bad type must obey:

$$\sum_{i\in\mathcal{O}_1^B} f_1^B\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}\right)=\sum_{i\in\mathcal{O}_2^B} f_i^B\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} f_i^B\tau^B_{i,\mathcal{O}_1^B}.$$

Rewriting it in terms of the likelihood ratios:

$$\sum_{i\in\mathcal{O}_1^B} \frac{f_i^B}{f_i^S} f_i^S\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}\right)=\sum_{i\in\mathcal{O}_2^B} \frac{f_i^B}{f_i^S} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} \frac{f_i^B}{f_i^S} f_i^S\tau^B_{i,\mathcal{O}_1^B}.$$

From the same reasoning as in the proof of Proposition 1, and recalling that there is at least one $i \in \mathcal{O}_1^B$ for which $\tau_{i,\mathcal{I}^B} > 0$, we obtain

$$\begin{aligned}
2\lambda\sum_{i\in\mathcal{O}_1^B} f_i^S\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}\right) &< \sum_{i\in\mathcal{O}_1^B} \frac{f_i^B}{f_i^S} f_i^S\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}\right),\\
&= \sum_{i\in\mathcal{O}_2^B} \frac{f_i^B}{f_i^S} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} \frac{f_i^B}{f_i^S} f_i^S\tau^B_{i,\mathcal{O}_1^B},\\
&\leq 2\lambda\left(\sum_{i\in\mathcal{O}_2^B} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} f_i^S\tau^S_{i,\mathcal{O}_1^B}\right).
\end{aligned}$$

But this contradicts (A.16), since it implies that

$$\sum_{i\in\mathcal{O}_1^B} f_i^S\left(\tau_{i,\mathcal{O}_2^B}+\tau_{i,\mathcal{I}^B}\right)<\left(\sum_{i\in\mathcal{O}_2^B} f_i^S\tau_{i,\mathcal{O}_1^B}+\sum_{i\in\mathcal{I}^B} f_i^S\tau^S_{i,\mathcal{O}_1^B}\right). \quad \square$$

Appendix B. Online appendix

In this section, we discuss and prove results about the applicability of our class of information policies, as well as the restrictions imposed by some of our assumptions in the paper. In the first section, we show that our rating systems with countably infinite rating sets can reproduce both the unbounded (full memory) and the bounded-recall environments. In the second section, we prove that our focus on deterministic hiring decisions for the customers is without loss of generality. In the third section, we show that there is still the same upper bound on the induced beliefs when we allow for calendar-time certainty.

Full memory and bounded recall

Our system can reproduce the full-memory and bounded-recall settings. Indeed, for any $t \in \mathbb{N}$ let $Z^t := \{H, L, Out\}^{t-1}$ be the set of past treatment sequences and $\mathcal{Z} = \cup_{t\in\mathbb{N}} Z^t$ be the set of all finite treatment sequences. Set $M = \mathcal{Z}$ and assign the following transition rule: for every $m \in M$ and every $y \in \{H, L, Out\}$, $\varphi^y_{m,m'} = 1$ if $m' = (m, y)$ (and zero otherwise). The notation $\varphi^y_{m,m'}$ represents the probability of the system moving from m to m' under y. Set also $\varphi_0 = \{\emptyset\}$. This is the full memory setting. Note that under full memory, every customer knows exactly the past history of actions and his position in time.

For the k-bounded recall setting, consider again $M = \mathcal{Z}$ and $\varphi_0 = \{\emptyset\}$, but assign now the following transition rule: (i) for every $m = z^t$ whose length is at most k and every $y \in \{H, L, Out\}$, $\varphi^y_{m,m'} = 1$ if $m' = (m, y)$; (ii) for every $m = z^t$ whose length is more than k and every y, $\varphi^y_{m,m'} = 1$ if $m' = (m_{-k}, y)$, m_{-k} being what is left of history $z^t = m$ after deleting the first $t - 1 - k$ entries of it. Then $\mathcal{M} = (M, \varphi, \varphi_0)$ represents a bounded-recall environment.

Since customers' only source of information comes from the rating system, it would be natural to consider they have an uniform prior over periods in the bounded recall setting, but with adaptations. Specifically, for periods $1, ..., k$, the length of the history associated with a rating would give the date away. However, for periods after k, customers would have an improper uniform prior and beliefs would be based on limit of the average of ratings distribution. The long-run frequency of observing a rating m under type $\omega \in \{S, B\}$ and a strategy profile $\sigma = (\alpha, \beta^H, \beta^L)$ would be

$$f^\omega_m := \lim_{t\to\infty} \left\{ \frac{1}{t-(k+1)} \sum_{s=k+1}^{t} P^s_{(\mathcal{M},\sigma)}(m|\omega) \right\}.$$

Random strategies for customers

In this subsection, we prove that our focus on deterministic hiring decisions was without loss of generality indeed.

Lemma 3. *For any rating system and strategy profile that induce an equilibrium with a randomized hiring decision, there exists another rating system and strategy profile inducing an equilibrium with deterministic hiring and leading to the same equilibrium payoffs.*

Proof. Consider a rating system $\mathcal{M} = (M, \varphi, \varphi_0)$ with n ratings and a strategy profile $(\alpha, \beta^H, \beta^L)$. Since customers are myopic, they hire in equilibrium in any rating satisfying equation (8) strictly and do not hire in any rating satisfying equation (7). Thus, a randomized hiring decision can only be an equilibrium if the system induces some rating k at which equation (8) holds with equality. So suppose that there exists one such rating and hiring in it occurs with probability $\alpha_k < 1$. Without loss of generality, suppose that the original system has the likelihood ratios ordered non-increasingly.

We can construct another rating system $\tilde{\mathcal{M}}$ with $n+1$ ratings and a deterministic strategy for the customers that induce the same steady-state probabilities in every rating different than k, for each type of expert. As such, the posterior beliefs in these ratings will be the same as the original one, so the hiring decisions in them will not change. Furthermore, the original strategy for the strategic expert remains optimal given the new system.

To do so, we first substitute rating k by ratings k_1 and k_2 and set all transition probabilities that neither enter or exit k to be the same as the ones from the original system: $\tilde{\varphi}^y_{ij} := \varphi^y_{ij}$ whenever $i, j \in \{1..., k-1, k+1, ...n\}$. Then we define the transitions to rating k_1 as

$$\tilde{\varphi}^y_{ik_1} := \varphi^y_{ik}(1-\alpha_k),$$

and transitions to rating k_2 as

$$\tilde{\varphi}^y_{ik_2} := \varphi^y_{ik}\alpha_k.$$

For transitions from k_1 and k_2 into $i \in \{1, ..., k-1, k+1, ..., n\}$, we set

$$\tilde{\varphi}^y_{k_1 i} = \tilde{\varphi}^y_{k_2 i} := \varphi^y_{ki}.$$

Finally, for transitions that keep the dynamics between k_1 and k_2 we set

$$\tilde{\varphi}^y_{k_1k_1} = (1-\alpha_k)\varphi^y_{kk}, \qquad \tilde{\varphi}^y_{k_2k_2} = \alpha_k\varphi^y_{kk},$$
$$\tilde{\varphi}^y_{k_2k_1} = (1-\alpha_k)\varphi^y_{kk}, \qquad \tilde{\varphi}^y_{k_1k_2} = \alpha_k\varphi^y_{kk}.$$

Note that all transition probabilities are well defined, for each $y \in \{H, L, Out\}$.

After constructing the rating system, define the new strategy profile $(\tilde{\alpha}, \tilde{\beta}^H, \tilde{\beta}^L)$ as follows. For customers: do not hire if $i \in \{1, ..., k_1\}$; hire if $i \in \{k_2, ..., n\}$. For the expert, for each $\theta \in \{H, L\}$: $\tilde{\beta}^\theta_i = \beta^\theta_i$ for all $i \in \{1, ..., k-1, k+1, ..., n\}$; $\tilde{\beta}^\theta_{k_1} = \tilde{\beta}^\theta_{k_2} = \beta^\theta_k$.

To simplify notation, let $\tau^{B,In}_{ki} := \varphi^H_{ki}$ and $\tau^{S,In}_{ki} := \gamma_k\varphi^H_{ki} + (1-\gamma_k)\varphi^L_{ki}$. Given the new rating system and the strategy profile, the conditional transitions $\tilde{\tau}^\omega_{ij}$ are the same as original ones τ^ω_{ij} whenever $i, j \in \{1, ..., k-1, k+1, ..., n\}$. The other transitions are as follows:

$$\tilde{\tau}^{\omega}_{ik_1} = \begin{cases} \varphi^{out}_{ik}(1-\alpha_k) & \text{if } i \in \{1,...,k-1\}, \\ \varphi^{Out}_{kk}(1-\alpha_k) & \text{if } i = k_1, \\ \tau^{\omega,In}_{kk}(1-\alpha_k) & \text{if } i = k_2, \\ \tau^{\omega}_{ik}(1-\alpha_k) & \text{if } i \in \{k+1,...n\}, \end{cases} \qquad \tilde{\tau}^{\omega}_{ik_2} = \begin{cases} \varphi^{out}_{ik}\alpha_k & \text{if } i \in \{1,...,k-1\}, \\ \varphi^{out}_{kk}\alpha_k & \text{if } i = k_1, \\ \tau^{\omega,In}_{kk}\alpha_k & \text{if } i = k_2, \\ \tau^{\omega}_{ik}\alpha_k & \text{if } i \in \{k+1,...n\}, \end{cases}$$

$$\tilde{\tau}^{\omega}_{k_1 i} = \left\{ \varphi^{out}_{ki} \quad \text{if } i \in \{1,...,k-1,k+1,...,n\}, \right. \qquad \tilde{\tau}^{\omega}_{k_2 i} = \left\{ \varphi^{\omega,In}_{ki} \quad \text{if } i \in \{1,...,k-1,k+1,...,n\}. \right.$$

Computing the equilibrium stationary distribution $\tilde{f}^{\omega}$, $\omega \in \{B,S\}$, under this new rating system and strategy profile gives:

$$\begin{aligned} \tilde{f}^{\omega}_i &= \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tilde{\tau}^{\omega}_{ji} + \tilde{f}^{\omega}_{k_1}\tilde{\tau}^{\omega}_{k_1 i} + \tilde{f}^{\omega}_{k_2}\tilde{\tau}^{\omega}_{k_2 i} \\ &= \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tau^{\omega}_{ji} + \tilde{f}^{\omega}_{k_1}\varphi^{Out}_{ki} + \tilde{f}^{\omega}_{k_2}\tau^{\omega,In}_{ki}, \end{aligned}$$

for all $i \neq k_1, k_2$. Denote by $\tilde{f}_k := \tilde{f}^{\omega}_{k_1} + \tilde{f}^{\omega}_{k_2}$ and by $\tilde{\tau}^{\omega}_{ki} := \frac{\tilde{f}^{\omega}_{k_1}\varphi^{Out}_{ki} + \tilde{f}^{\omega}_{k_2}\tau^{\omega,In}_{ki}}{\tilde{f}^{\omega}_{k_1}+\tilde{f}^{\omega}_{k_2}}$. Straightforward manipulations of the equation above leads to:

$$\tilde{f}^{\omega}_i = \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tau^{\omega}_{ji} + \tilde{f}^{\omega}_k \tilde{\tau}^{\omega}_{ki}. \tag{B.1}$$

For rating k_1, we have

$$\begin{aligned} \tilde{f}^{\omega}_{k_1} &= \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tilde{\tau}^{\omega}_{jk_1} + \tilde{f}^{\omega}_{k_1}\tilde{\tau}^{\omega}_{k_1k_1} + \tilde{f}^{\omega}_{k_2}\tilde{\tau}^{\omega}_{k_2k_1}, \\ &= \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tau^{\omega}_{jk}(1-\alpha_k) + \tilde{f}^{\omega}_{k_1}\varphi^{Out}_{kk}(1-\alpha_k) + \tilde{f}^{\omega}_{k_2}\tau^{\omega,In}_{kk}(1-\alpha_k), \end{aligned} \tag{B.2}$$

and for rating k_2 we have

$$\begin{aligned} \tilde{f}^{\omega}_{k_2} &= \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tilde{\tau}^{\omega}_{jk_2} + \tilde{f}^{\omega}_{k_1}\tilde{\tau}^{\omega}_{k_1k_2} + \tilde{f}^{\omega}_{k_2}\tilde{\tau}^{\omega}_{k_2k_2}, \\ &= \sum_{j\in\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \varphi^{\omega}_{jk}\alpha_k + \tilde{f}^{\omega}_{k_1}\varphi^{Out}_{kk}\alpha_k + \tilde{f}^{\omega}_{k_2}\tau^{\omega,In}_{kk}\alpha_k. \end{aligned} \tag{B.3}$$

Summing equations (B.2) and (B.3), we get

$$\tilde{f}^{\omega}_k = \sum_{j=\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tau^{\omega}_{jk} + \tilde{f}^{\omega}_{k_1}\varphi^{Out}_{kk} + \tilde{f}^{\omega}_{k_2}\tau^{\omega,In}_{kk}.$$

Define $\tilde{\tau}^{\omega}_{kk} := \frac{\tilde{f}^{\omega}_{k_1}\varphi^{Out}_{kk} + \tilde{f}^{\omega}_{k_2}\tau^{\omega,In}_{kk}}{\tilde{f}^{\omega}_{k_1}+\tilde{f}^{\omega}_{k_2}}$. Then,

$$\tilde{f}^{\omega}_k = \sum_{j=\{1,...,k-1,k+1,...,n\}} \tilde{f}^{\omega}_j \tau^{\omega}_{jk} + \tilde{f}^{\omega}_k \tilde{\tau}^{\omega}_{kk}. \tag{B.4}$$

Note as well from equations (B.2) and (B.3) that $\tilde{f}^{\omega}_{k_1} = (1-\alpha_k)\tilde{f}^{\omega}_k$ and $\tilde{f}^{\omega}_{k_2} = \alpha_k \tilde{f}^{\omega}_k$, respectively. Therefore, $\tilde{\tau}^{\omega}_{kj} = \tau^{\omega}_{kj} = \alpha_k \tau^{\omega,In}_{kj} + (1-\alpha_k)\varphi^{Out}_{kj}$, for $j \in M$. Substituting this identity into equations (B.1) and (B.4), we obtain

$$\tilde{f}^{\omega}_i = \sum_{j\in M} \tilde{f}^{\omega}_j \tau^{\omega}_{ji} \;\; \forall \; i \in M.$$

This shows that the new distribution $\tilde{f}^{\omega}$ is also a stationary distribution for the conditional stochastic matrix T^{ω}. Finally, define the following initial distribution: $\tilde{\varphi}_{0,i} := \varphi_{0,i}$ for $j \in \{1,...,k-1,k+1,...,n\}$; $\tilde{\varphi}_{0,k_1} = (1-\alpha_k)\varphi_{0,k}$ and $\tilde{\varphi}_{0,k_2} = \alpha_k\varphi_{0,k}$. Define as well $\tilde{\varphi}_{0,k} := \tilde{\varphi}_{0,k_1} + \tilde{\varphi}_{0,k_2}$. Then, under $(\tilde{\varphi}_{0,i})_{i\in M}$ and the strategies $(\tilde{\alpha}_i)_{i\in M}$, $(\tilde{\beta}^H_i, \tilde{\beta}^L_i)_{i\in M}$ converges to $(f^{\omega})_{i\in M}$.

It remains to check whether the new strategy profile is an equilibrium. Because $\tilde{f}^{\omega}_i = f^{\omega}_i$ for all $i \in \{1,...,k-1,k+1,...,n\}$ and each $\omega \in \{B,S\}$, the posterior beliefs in such ratings will be identical. In rating k_1, the posterior belief will be

$$\tilde{\rho}_{k_1} = \frac{\rho\alpha_k f^B_k}{\rho\alpha_k f^B_k + (1-\rho)\alpha_k f^S_k} = \rho_k,$$

which is exactly the belief under the original system in rating k. Because $\tilde{\beta}_{k_1} = \beta_k$, it is optimal for customers to have $\tilde{\alpha}_{k_1} = 0$. Likewise, in rating k_2, $\tilde{\rho}_{k_2} = \rho_k$, and it is optimal for customers as well to have $\tilde{\alpha}_{k_2} = 1$.

To see that the original strategy profile for the expert is still optimal and the equilibrium payoff is the same, start by defining $\tilde{V}_k := (1-\alpha_k)\tilde{V}_{k_1} + \alpha_k \tilde{V}_{k_2}$. Then every continuation value at $i \in M$ can be written as

$$\tilde{V}_i = (1-\delta)\alpha_i\,[\beta_i u - (1-\beta_i)w] + \delta \sum_{j \in M} \tau_{ij}^S \tilde{V}_j.$$

The above system of equations also shows that the new continuation values over M remain the same as the values under the original system and equilibrium. As such, the expert's strategy profile remains optimal in all $i \in \{1, \ldots, k-1, k+1, \ldots, n\}$. In rating k_1, he is not hired in equilibrium and in k_2 if he plays $\beta_{k_2}^{\theta}$ with minimum probabilities to justify hiring, then

$$\begin{aligned}
\tilde{V}_{k_2}^{\theta} &= (1-\delta)\left[\beta_{k_2}^{\theta} u - (1-\beta_{k_2}^{\theta})w\right] + \delta \sum_{m' \in \tilde{M}} \left\{\beta_{k_2}^{\theta}\varphi_{k_2 m'}^{y=\theta} + (1-\beta_{k_2}^{\theta})\varphi_{k_2 m'}^{y\in\Theta\setminus\{\theta\}}\right\}\tilde{V}_{m'},\\
&= (1-\delta)\left[\beta_{k_2}^{\theta} u - (1-\beta_{k_2}^{\theta})w\right] + \delta \sum_{m' \in \{1,\ldots,k-1,k+1,\ldots,n\}} \left\{\beta_{k_2}^{\theta}\varphi_{km'}^{y=\theta} + (1-\beta_{k_2}^{\theta})\varphi_{km'}^{y\in\Theta\setminus\{\theta\}}\right\}\tilde{V}_{m'}+\\
&+ \delta\{\beta_{k_2}^{\theta}\varphi_{kk}^{y=\theta} + (1-\beta_{k_2}^{\theta})\varphi_{kk}^{y\in\Theta\setminus\{\theta\}}\}\left[(1-\alpha_k)\tilde{V}_{k_1} + \alpha_k\tilde{V}_{k_2}\right],\\
&= (1-\delta)\left[\beta_{k_2}^{\theta} u - (1-\beta_{k_2}^{\theta})w\right] + \delta \sum_{m \in M} \left\{\beta_{k_2}^{\theta}\varphi_{k_2 m}^{y=\theta} + (1-\beta_{k_2}^{\theta})\varphi_{k_2 m}^{y\in\Theta\setminus\{\theta\}}\right\}V_m.
\end{aligned}$$

Therefore, $\beta_{k_2}^{\theta} = \beta_k^{\theta}$ reaches the maximum continuation value V_k at k_2. If the expert deviates to a probability that triggers customers not to hire at k_2, then he obtains

$$\delta \sum_{m \in M} \varphi_{km}^{Out} V_m.$$

Since in the original system it is optimal for the strategic expert to play β_k^{θ} and this is such that there exists positive hiring in k, it must be that $\beta_{k_2}^{\theta} = \beta_k^{\theta}$ is optimal in the new system. □

Calendar time awareness

In the third section of the online appendix, we prove that, even if customers are aware of calendar time, there is the same bound on what rating systems can achieve. Suppose first that $\rho > \rho^*$. Then there is a cold-start problem. Customer 1 will not hire, no information will be generated, so customer 2 does not hire, regardless of the rating system in place. Subsequent customers do not hire either since no new information is being revealed and the market breaks down. Thus, knowing the calendar time in this environment will not permit a functioning market.

Now, let us consider the case in which $\rho \le \rho^*$, so that the first customer hires if the expert's strategy is to perform the correct treatment with sufficiently high probability. First, consider the optimal rating system for the expert. It suffices to construct a rating system that does not display any information to future consumers. In this case, all consumers hire, and the expert's optimal strategy is to choose the correct treatment at all periods. This is his first best and can be achieved under these parameters irrespective of the assumption that players know the calendar time.

Let us now think about the optimal rating system for the customers (again, when $\rho \le \rho^*$). The next proposition proves that the bound on the beliefs that can be induced in equilibria is $\bar{\rho}$.

Lemma 4. *Assume that customers are aware of calendar time. Consider any rating system and equilibrium in which there is a period t such the maximum posterior belief at t is at most $\bar{\rho}$. Then the highest posterior belief in period $t+1$ is also at most $\bar{\rho}$.*

Proof. Let i be an arbitrary rating at period $t+1$. Since consumers are aware of calendar time, each pair rating and time define an information set. They compute each type's probability of the system reaching rating i at period t, denoted by $f_{j,t}^B$, as the sum of the probabilities of reaching every node in that information set given type B, and similarly for type S.

Define also $\tau_{ji}^{B,t}$ to be the equilibrium probability with which the rating system moves from rating j to rating i at time period t if the type is B. Since type B always choose treatment t_H, this probability is simply the transition probability, that is

$$\tau_{ji}^{B,t} = \varphi_{ji}^H.$$

In the case of the strategic type, $\tau_{ji}^{S,t}$ depends on both φ_{ji}^H and φ_{ji}^L and the equilibrium strategy of the strategic type: $\beta_{j,t}^H$ and $\beta_{j,t}^L$. Thus, we may write

$$\begin{aligned}
\tau_{ji}^{S,t} &= \frac{1}{2}\left[(1-\beta_{j,t}^L)\varphi_{ji}^H + \beta_{j,t}^L\varphi_{ji}^L\right] + \frac{1}{2}\left[\beta_{j,t}^H\varphi_{ji}^H + (1-\beta_{j,t}^H)\varphi_{ji}^L\right]\\
&= \gamma_{j,t}\varphi_{ji}^H + (1-\gamma_{j,t})\varphi_{ji}^L,
\end{aligned}$$

where $\gamma_{j,t} := \frac{1}{2}\beta^H_{j,t} + \frac{1}{2}(1-\beta^L_{j,t})$ is the expert's probability of providing the high-cost treatment. We can write the probabilities of reaching rating i in period $t+1$ as

$$f^B_{i,t+1} = \sum_{j\in M} f^B_{j,t}\tau^{B,t}_{ji}, \tag{B.5}$$

$$f^S_{i,t+1} = \sum_{j\in M} f^S_{j,t}\tau^{S,t}_{ji}. \tag{B.6}$$

Thus, the posterior belief about expert being bad at $t+1$ will be

$$\rho_{i,t+1} = \frac{\rho f^B_{i,t+1}}{\rho f^B_{i,t+1} + (1-\rho) f^S_{i,t+1}}. \tag{B.7}$$

Define $f_{j,t} := \rho f^B_{j,t} + (1-\rho) f^S_{j,t}$. Using such definition and plugging equations (B.5) and (B.6) in equation (B.7) above, we get

$$\rho_{i,t+1} = \frac{\sum_{j\in M} f_{j,t}\left[\frac{\rho f^B_{j,t}}{f_{j,t}}\right]\tau^{B,t}_{ji}}{\sum_{j\in M} f_{j,t}\left[\frac{\rho f^B_{j,t}}{f_{j,t}}\right]\tau^{B,t}_{ji} + \sum_{j\in M} f_{j,t}\left[\frac{(1-\rho) f^S_{j,t}}{f_{j,t}}\right]\tau^{S,t}_{ji}}.$$

Using the definition of beliefs as in equation (B.7), but for period t,

$$\rho_{i,t+1} = \frac{\sum_{j\in M} f_{j,t}\rho_{j,t}\tau^{B,t}_{ji}}{\sum_{j\in M} f_{j,t}\left[\rho_{j,t}\tau^{B,t}_{ji} + (1-\rho_{j,t})\tau^{S,t}_{ji}\right]}.$$

Let $\mathcal{O}_t$ refer to non-hiring ratings and $\mathcal{I}_t$ to hiring ratings at t. Then,

$$\begin{aligned}\rho_{i,t+1} &= \frac{\sum_{j\in\mathcal{O}_t} f_{j,t}\rho_{j,t}\varphi^{Out}_{ji} + \sum_{j\in\mathcal{I}_t} f_{j,t}\rho_{j,t}\tau^{B,t}_{ji}}{\sum_{j\in\mathcal{O}_t} f_{j,t}\varphi^{Out}_{ji} + \sum_{j\in\mathcal{I}_t} f_{j,t}\left[\rho_{j,t}\tau^{B,t}_{ji} + (1-\rho_{j,t})\tau^{S,t}_{ji}\right]}, \\ &= \frac{\sum_{j\in\mathcal{O}_t} f_{j,t}\rho_{j,t}\varphi^{Out}_{ji} + \sum_{j\in\mathcal{I}_t} f_{j,t}\left[\rho_{j,t}\tau^{B,t}_{ji} + (1-\rho_{j,t})\tau^{S,t}_{ji}\right]\left\{\frac{\rho_{j,t}\tau^{B,t}_{ji}}{\rho_{j,t}\tau^{B,t}_{ji}+(1-\rho_{j,t})\tau^{S,t}_{ji}}\right\}}{\sum_{j\in\mathcal{O}_t} f_{j,t}\varphi^{Out}_{j,i} + \sum_{j\in\mathcal{I}_t} f_{j,t}\left[\rho_{j,t}\tau^{B,t}_{ji} + (1-\rho_{j,t})\tau^{S,t}_{ji}\right]}.\end{aligned} \tag{B.8}$$

Before we proceed, note that in period t the customer was only willing to hire if in equilibrium her payoff of hiring was greater than his payoff of not hiring, that is:

$$\rho_{j,t}\left(\frac{u-w}{2}\right) + (1-\rho_{j,t})\left[\frac{1}{2}\left(-(1-\beta^L_{j,t})w + \beta^L_{j,t}u\right) + \frac{1}{2}\left(\beta^H_{j,t}u - (1-\beta^H_{j,t})w\right)\right] \geq 0$$

Simplifying, $(u-w) + (1-\rho_{j,t})\left[\beta^H_{j,t} - (1-\beta^L_{j,t})\right](u+w) \geq 0$. Finally,

$$\beta^H_{j,t} \geq \frac{w-u}{w+u}\frac{1}{(1-\rho_{j,t})} + 1 - \beta^L_{j,t}.$$

Thus, in equilibrium, hiring requires that the strategic type performs the correct treatment with a minimum probability when the problem is H: $\beta^H_{j,t} \geq \frac{w-u}{w+u}\frac{1}{(1-\rho_{j,t})}$. Therefore, the transition probability $\tau^{S,t}_{ji}$ is bounded:

$$\tau^{S,t}_{ji} \geq \frac{1}{2}\{\varphi^H_{ji}\beta^H_{j,t}\} = \frac{1}{2}\left\{\varphi^H_{ji}\frac{w-u}{w+u}\frac{1}{(1-\rho_{j,t})}\right\}.$$

We can substitute $(1-\rho^*) = \frac{w-u}{w+u}$ and we can find a bound on the term in curly brackets $\frac{\rho_{j,t}\tau^{B,t}_{ji}}{\rho_{j,t}\tau^{B,t}_{ji}+(1-\rho_{j,t})\tau^{S,t}_{ji}}$ in equation (B.7). That is,

$$\begin{aligned}\frac{\rho_{j,t}\tau^{B,t}_{ji}}{\rho_{j,t}\tau^{B,t}_{ji} + (1-\rho_{j,t})\tau^{S,t}_{ji}} &\leq \frac{\rho_{j,t}\varphi^H_{ji}}{\rho_{j,t}\tau^{B,t}_{ji} + (1-\rho_{j,t})\frac{1}{2}\{\varphi^H_{ji}(1-\rho^*)\frac{1}{(1-\rho_{j,t})}\}}, \\ &= \frac{2\rho_{j,t}}{2\rho_{j,t} + (1-\rho^*)},\end{aligned}$$

$$\leq \frac{2\rho^*}{2\rho^* + (1-\rho^*)},$$

where the last inequality comes from the fact that the expert was hired, and thus $\rho_{j,t} \leq \rho^*$. Therefore,

$$\frac{\rho_{j,t}\tau_{ji}^{B,t}}{\rho_{j,t}\tau_{ji}^{B,t} + (1-\rho_{j,t})\tau_{ji}^{S,t}} \leq \frac{2\rho^*}{1+\rho^*} = \bar{\rho}.$$

Substituting it into equation (B.8), we conclude that

$$\rho_{i,t+1} \leq \bar{\rho} \left\{ \frac{\sum_{j\in\mathcal{O}_t} f_{j,t}\varphi_{ji}^{Out} + \sum_{j\in\mathcal{I}_t} f_{j,t}\left[\rho_{j,t}\tau_{ji}^{B,t} + (1-\rho_{j,t})\tau_{ji}^{S,t}\right]}{\sum_{j\in\mathcal{O}_t} f_{j,t}\varphi_{ji}^{Out} + \sum_{j\in\mathcal{I}_t} f_{j,t}\left[\rho_{j,t}\tau_{ji}^{B,t} + (1-\rho_{j,t})\tau_{ji}^{S,t}\right]} \right\} = \bar{\rho}. \quad \square$$

The above lemma implies that, if $\rho \leq \rho^*$ (consequently, $\rho < \bar{\rho}$), then in every period the maximum posterior belief is at most $\bar{\rho}$. So, with calendar time awareness, the constrained on the belief spread still remains. If we consider the perspective of a social planner that seeks to maximize the discounted average payoff of myopic customers, then the next proposition proves that the maximum equilibrium payoff is also constrained, and by the same constraint as the one in Proposition 1 in the text.

Proposition 6. *Assume that customers are aware of calendar time. At every period $t \in \mathbb{N}$, for every prior belief $\rho \in (0, \bar{\rho})$, the social planner's maximum equilibrium payoff is bounded above:*

$$v_t \leq u\left[\frac{\bar{\rho}-\rho}{\bar{\rho}}\right].$$

Proof. Consider any time $t \in \mathbb{N}$ and rating $i \in M$. A social planner's value functions, conditional on interacting with each type of expert, are as follows.

$$v_{t,i}^B = (1-\delta)\alpha_{t,i}\left(\frac{u-w}{2}\right) + \delta\sum_{j\in M}\tau_{ij}^{B,t}v_{t+1,j}^B \tag{B.9}$$

$$v_{t,i}^S = (1-\delta)\alpha_{t,i}\left[\beta_{t,i}u - (1-\beta_{t,i})w\right] + \delta\sum_{j\in M}\tau_{ij}^{S,t}v_{t+1,j}^S. \tag{B.10}$$

If i is a hiring rating, then the unconditional value function is bounded above by the highest payoff customers can get in a one-shot interaction:

$$\begin{aligned}
v_{t,i} &= \rho_{i,t}v_{i,t}^B + (1-\rho_{i,t})v_{i,t}^S, \\
&= (1-\delta)\nu_{i,t} + \delta\sum_{j\in M}\left[\rho_{i,t}\tau_{ij}^{B,t}v_{j,t+1}^B + (1-\rho_{i,t})\tau_{ij}^{S,t}v_{j,t+1}^S\right], \\
&= (1-\delta)\nu_{i,t} + \delta\sum_{j\in M}\tilde{\tau}_{i,j}^t\left[\tilde{\rho}_{j,t+1}(i)v_{j,t+1}^B + (1-\tilde{\rho}_{j,t+1}(i))v_{j,t+1}^S\right], \\
&\leq (1-\delta)u + \delta\sum_{j\in M}\tilde{\tau}_{ij}^t u, \\
&= u,
\end{aligned}$$

where $\tilde{\tau}_{ij}^t := \rho_{i,t}\tau_{ij}^{B,t} + (1-\rho_{i,t})\tau_{ij}^{S,t}$ and $\tilde{\rho}_{j,t+1}(i) := \frac{\rho_{i,t}\tau_{ij}^{B,t}}{\tilde{\tau}_{ij}^t}$. The inequality follows from $\nu_{i,t} \leq u$ as well as $\max\{v_{j,t+1}^B, v_{j,t+1}^S\} \leq u$ for any j. Therefore, if $\mathcal{I}_t$ represents the set of all hiring ratings in period t, then

$$v_t = \sum_{i\in\mathcal{I}_t} f_{i,t}v_{i,t} \leq u\sum_{i\in\mathcal{I}_t} f_{i,t}.$$

Following the same arguments from Proposition 2, noting that Bayes plausibility holds even with calendar certainty - that is, $\sum_{i\in M} f_{i,t}\rho_{i,t} = \rho$, and from Lemma 4, we can conclude that $v_t \leq u\left[\frac{\bar{\rho}-\rho}{\rho}\right]$. $\square$

References

Aumann, R.J., Maschler, M., Stearns, R.E., 1995. Repeated Games with Incomplete Information. MIT Press, Cambridge.
Bertsekas, D., 2015. Dynamic Programming and Optimal Control: Volume II, 4 ed. Athena Scientific.
Bright Local, 2018. Local consumer review survey 2018. https://www.brightlocal.com/research/local-consumer-review-survey-2018/. (Accessed 22 March 2021).
Bright Local, 2019. Local consumer review survey 2019. https://www.brightlocal.com/research/local-consumer-review-survey-2019/. (Accessed 22 March 2021).
Compte, O., Postlewaite, A., 2015. Plausible cooperation. Games Econ. Behav. 91, 45–59.
Cripps, M.W., Mailath, G.J., Samuelson, L., 2004. Imperfect monitoring and impermanent reputations. Econometrica 72, 407–432.
Dellarocas, C., Wood, C.A., 2008. The sound of silence in online feedback: estimating trading risks in the presence of reporting bias. Manag. Sci. 54, 460–476.
Ekmekci, M., 2011. Sustainable reputations with rating systems. J. Econ. Theory 146, 479–503.
Ely, J.C., Välimäki, J., 2003. Bad reputation. Q. J. Econ. 118, 785–814.
Filippas, A., Horton, J.J., Golden, J., 2018. Reputation inflation. In: Proceedings of the 2018 ACM Conference on Economics and Computation, pp. 483–484.
Fudenberg, D., Levine, D.K., 1989. Reputation and equilibrium selection in games with a patient player. Econometrica 57, 759–778.
Fudenberg, D., Levine, D.K., 1992. Maintaining a reputation when strategies are imperfectly observed. Rev. Econ. Stud. 59, 561–579.
Goldman, E., 2010. The regulation of reputational information. In: Szoka, B., Marcus, A. (Eds.), The Next Digital Age: Essays on the Future of the Internet. TechFreedom. Washington, D.C., pp. 293–304. Chapter 4.
Hellman, M.E., Cover, T.M., 1970. Learning with finite memory. Ann. Math. Stat., 765–782.
Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. Am. Econ. Rev. 101, 2590–2615.
Lillethun, E., 2017. Optimal information design for reputation building. Working paper. Colgate University.
Lorecchio, C., Monte, D., 2023. Dynamic information design under constrained communication rules. Am. Econ. J. Microecon. 15, 359–398.
Mailath, G.J., Samuelson, L., 2006. Repeated Games and Reputations: Long-Run Relationships. Oxford University Press, New York, NY.
Monte, D., 2013. Bounded memory and permanent reputations. J. Math. Econ. 49, 345–354.
Monte, D., 2014. Learning with bounded memory in games. Games Econ. Behav. 87, 204–223.
Nosko, C., Tadelis, S., 2015. The limits of reputation in platform markets: an empirical analysis and field experiment. Working paper. National Bureau of Economic Research.
Sperisen, B., 2018. Bad reputation under bounded and fading memory. Econ. Inq. 56, 138–157.
Stokey, N., Lucas, R., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, Massachusetts.
Vong, A., 2021. Firm certification. Working paper. Yale University.
Wilson, A., 2014. Bounded memory and biases in information processing. Econometrica 82, 2257–2294.
Zervas, G., Proserpio, D., Byers, J.W., 2021. A first look at online reputation on airbnb, where every stay is above average. Mark. Lett. 32, 1–16.