Rating Protocols in Online Communities

Yu Zhang, Jaeok Park [12], Mihaela van der Schaar [3]

Abstract—*Sustaining cooperation among self-interested agents is critical for the proliferation of emerging online communities. Providing incentives for cooperation in online communities is particularly challenging because of their unique features: a large population of anonymous agents having asymmetric interests and dynamically joining and leaving the community; operation errors; and agents trying to whitewash when they have a low standing in the community. In this paper, we take these features into consideration and propose a framework for designing and analyzing a class of incentive schemes based on rating protocols, which consist of a rating scheme and a recommended strategy. We first define the concept of sustainable rating protocols under which every agent has incentives to follow the recommended strategy given the deployed rating scheme. We then formulate the problem of designing an optimal rating protocol, which selects the protocol that maximizes the overall social welfare among all sustainable rating protocols. Using the proposed framework, we study the structure of optimal rating protocols and explore the impact of one-sided rating, punishment lengths and whitewashing on optimal rating protocols. Our results show that optimal rating protocols are capable of sustaining cooperation, with the amount of cooperation varying depending on the community characteristics.*

Categories and Subject Descriptors: **H.4.0 [Information System Applications]**: General; **J.4 [Social and Behavioral Sciences]**: Economics

General Terms: Design, Economics, Human Factors

Additional Keywords and Phrases: Incentive schemes, online communities, rating schemes, recommended strategy, whitewashing.

I. INTRODUCTION

Recent developments in technology have expanded the boundaries of communities in which individuals interact with each other. For example, nowadays individuals can obtain valuable information or content from remotely located individuals in an online community formed through online networking services [1]-[7]. However, a large population and the anonymity of individuals in such an online community make it difficult to sustain cooperative behavior among self-interested individuals [8][9]. For example, it has been reported that "free-riding" is widely observed in peer-to-peer networks [10][11]. Hence, incentive schemes are needed to cultivate cooperative behavior in online communities.

A variety of incentive schemes have been explored to induce cooperation in such online communities. The most popular incentives are based on pricing schemes and differential service provision. Pricing schemes use payments to reward and punish individuals for their behavior, which in principle can induce self-interested individuals to cooperate with each other to attain the social optimum by internalizing their external effects (see, for example, [12][13]). However, pricing schemes often require complex accounting infrastructures, which introduce substantial communication and computation overhead [14]. Hence, it is impractical for pricing schemes and the corresponding infrastructure to be implemented in open, distributed communities, e.g. peer-to-peer networks [17], mobile networks [2], etc., where agents have limited computing and communication capabilities. Moreover, the operators of online communities may be reluctant to adopt a pricing scheme when pricing discourages individuals' participation in community activities. Finally and most importantly, the services being exchanged in online communities are often not real goods but rather solutions to small tasks or provisions of small amounts of resources

---

[1] Jaeok Park is with School of Economics, Yonsei University, Korea. e-mail: jaeok.park@yonsei.ac.kr

[2] This work was done when Jaeok Park was a postdoctoral scholar at Electrical Engineering Department, University of California, Los Angeles, USA.

[3] Yu Zhang and Mihaela van der Schaar are with Electrical Engineering Department, University of California, Los Angeles, USA. e-mail: yuzhang@ucla.edu, mihaela@ee.ucla.edu.

which are *difficult to price*. For example, in the online question and answer forums such as Yahoo! Answers [1], the exchanged service represents a small "favor" in answering the questions posted by other users. The difficulty in pricing such "small" services in these applications prevents pricing scheme to be effective.

Differential service schemes, on the other hand, reward and punish individuals by providing differential services depending on their behavior instead of using monetary rewards [15]-[30]. Differential services can be provided by community operators or by community members. Community operators can treat individuals differentially (for example, by varying the quality or scope of services) based on the information about the behavior of individuals. Incentive provision by a central entity can offer a robust method to sustain cooperation [15]. However, such an approach is impractical in a large community because the burden of a central entity to monitor individuals' behavior and provide differential services for them becomes prohibitively heavy as the population size grows. Alternatively, more distributed incentive schemes exist where community members monitor the behavior of each other and provide differential services based on their observations [16]-[30]. Such incentive schemes are based on the principle of *reciprocity* and can be classified into *personal reciprocation* (or direct reciprocity) [16]-[18] and *social reciprocation* (or indirect reciprocity) [19]-[30]. In personal reciprocation schemes, individuals can identify each other, and behavior toward an individual is based on their personal experience with that individual. *Personal reciprocation* is effective in sustaining cooperation in a small community where individuals can identify each other and interact frequently with fixed opponents, but it loses its power in a large community where individuals have asymmetric interests and can freely and frequently change the opponents they interact with [26]. In *social reciprocation* schemes, individuals obtain some information about other individuals (for example, rating) and decide their actions toward an individual based on this available information. Hence, an individual can be rewarded or punished by other individuals in the online community who have not had past interactions with him [26][27]. Therefore, social reciprocation has a potential to form a basis of successful incentive schemes for online communities. As such, this paper is devoted to the study of incentive schemes based on social reciprocation.

Sustaining cooperation using social reciprocation has been investigated in the literature using the framework of anonymous random matching games, in which each individual is repeatedly matched with different partners over time for service exchange and tries to maximize his discounted long-term utility. To implement social reciprocation, it is important for the community to share enough information about past interactions such that the community members know how to reward or punish others. This existing literature makes different assumptions on the information revealed to community members about other members. In [28] each community member observes the entire history of the past plays of his current partner. In [29][30], community members are informed about the outcomes of the matches in which they have been directly involved. Rating protocols have been proposed in [26] and [27], where each community member is attached a rating score indicating his social status, which takes a value from a finite set and records his past plays, and community members with different rating scores are treated differently by other individuals they interact with. For online communities, maintaining direct records of individuals' past plays which are used in [28]-[30] are not appropriate, because the communication and storage cost for revealing the entire history of the past plays of an individual grow unbounded with time. Since the use of rating score as a summary record requires significantly less amount of information to be maintained, we will design incentive schemes based on rating protocols for online communities. A rating-based incentive scheme can be easily implemented in online communities that deploy entities (e.g., a tracker in P2P networks [17], or a web portal in web-based applications [1]) that can collect, process, and deliver information about individuals' play history to generate rating scores.

Cooperation among community members can be sustained in all the above works on anonymous random matching games. However, all of them have focused on obtaining the Folk Theorem by characterizing the set of equilibrium payoffs that can be achieved when the discount factor of individuals is sufficiently close to 1. Our work, on the contrary, addresses the problem of designing a rating-based incentive scheme given a discount factor and other parameters arising from practical considerations, which are not fully considered in the existing literature on anonymous random matching games. Specifically, our work takes into account the following features of online communities:

- *Asymmetry of interests*. As an example, consider a community where individuals with different areas of expertise share knowledge with each other. It will be rarely the case that a pair of individuals has a mutual interest in the expertise of each other simultaneously. We allow the possibility of asymmetric interests by modeling the interaction between a pair of individuals as a gift-giving game, instead of a prisoner's dilemma game which assumes mutual interests between a pair of individuals [17][26][27][28]. [4]
- *Report errors*. In an incentive scheme based on a rating protocol, it is possible that the rating score (or label) of a specific individual is updated incorrectly because of errors in the report of his partners (i.e. other individuals he interacts with). Our model incorporates the possibility of report errors, which allows us to analyze its impact on design and performance, whereas most existing works on rating schemes (e.g. [26][27]) adopt an idealized assumption that rating scores are always updated correctly.
- *Dynamic change in the population*. The members of a community change over time as individuals gain or lose interest in the services provided by community members. We model this feature by having a fraction of individuals leave and join the community in every period. This allows us to study the impact of population turnover on design and performance.
- *Whitewashing*. Whitewashing refers to the behavior of an individual creating multiple identities by repeatedly entering to an anonymous online community. In an online community, individuals with bad rating scores may attempt to whitewash their rating scores by leaving and rejoining the community as new members to avoid the punishments imposed by the system upon their old identities [17]. We consider this possibility and study the design of "whitewash-proof" rating protocols and their performance.

Note that our model and analysis also differ significantly from most existing works on reputation systems [23]-[25]. First, the models in [23]-[25] assume that individuals assume fixed roles in the community (i.e. seller or buyer), which is common in applications where the groups of sellers and buyers are separated and usually do not overlap [21]. Nevertheless, in online communities such as P2P networks, online labor markets, etc., each agent can be both the provider and the receiver of services. Second, the reputation systems in [23]-[25] rely on differential

[4] In this work, we specifically use the "asymmetry of interests" to refer to the fact that two individuals in one stage game do not have mutual interests in the resources possessed by each other simultaneously. The investigation on other dimensions of individual asymmetry (e.g. heterogeneous individual characteristics) serves as an interesting extension of this work.

pricing schemes to incentivize sellers to cooperate. We have already mentioned that the services being exchanged and shared in online communities are difficult to price, thereby preventing such pricing-based reputation systems to be effectively deployed. Finally and most importantly, [23]-[25] consider the repeated game between a unique long-lived seller and many short-lived buyers, and the design principle there is to maximize the expected (discounted) long-term utility of the individual long-lived seller when the discount factor of the seller and the payment-to-cost ratio are sufficiently large. In contrast, we consider in this paper the interplay among a large number of long-lived individuals, and aim to maximize the social welfare (i.e. the sum utility) of the entire community for arbitrary discount factors and payment-to-cost ratios, which makes the designs in [23]-[25] inapplicable here. A more detailed and in-depth comparison between our work and [23]-[25] is provided in Section VII. The differences between our work and the existing literature on social reciprocity are summarized in the following table in order to highlight our contribution and novelty.

TABLE 1. Comparison between the existing literature and our work.

	[28]-[30][35]	[26][27]	[23]-[25]	Our work
Incentive device	Differential services	Differential services	Monetary rewards	Differential services
Asymmetry of interests	N/A	No	No	Yes
Report errors	N/A	No	Yes	Yes
Information requirement	Entire history of stage game outcomes	Individual rating	Individual rating	Individual rating
Discount factor	Sufficiently close to 1	Sufficiently close to 1	Sufficiently large	Arbitrary
Number of long-lived players	Multiple	Multiple	One	Multiple
Protocol design	No	No	Yes	Yes
Optimization criterion	Individual long-term utility	Individual long-term utility	Individual long-term utility	Sum utility of all players

The remainder of this paper is organized as follows. In Section II, we describe the repeated anonymous matching game and incentive schemes based on a rating protocol. In Section III, we formulate the problem of designing an optimal rating protocol. In Section IV, we provide analytical results about optimal rating protocols. In Section V, we extend our model to address the impacts of variable punishment lengths, whitewashing possibility, and one-sided rating. We provide simulation results in Section VI, and conclude the paper in Section VII.

II. MODEL

A. Repeated Matching Game

We consider a community where each member, or agent, can offer a valuable service to other agents. Examples of services are expert knowledge, customer reviews, job information, multimedia files, storage space, and computing power. We consider an infinite-horizon discrete time model with a continuum of agents [17]. [5] In a

[5] The continuum population model is commonly adopted in the analysis for large-scale dynamic networks, e.g. peer-to-peer systems [3][17], grid networks [4], social sharing websites [1][5], etc. Also, it has been shown in our technical report [22] that the continuum model can significantly reduce the complexity of designing optimal rating protocols in the anonymous random matching game while incurring small efficiency loss compared to the case where a finite population model is employed.

period, each agent generates a service request [37], which is sent to another agent that can provide the requested service. [6] We model the request generation and agent selection process using *uniform random matching*: each agent receives exactly one request in every period and each agent is equally likely to receive the request of an agent, and the matching is independent across periods. [7] Such model well approximates the matching process between agents in large-scale online communities. For example, in a mobile relay network [2] where agents (e.g. mobile devices) within a certain area are able to relay traffic for each other through unlicensed spectrum (e.g. WLAN) to the destination (e.g. nearby cellular base stations), the relay node that each mobile agent encounters at each moment could be approximately assumed to be random, since this mobile agent is moving around the area randomly over time.

In a pair of matched agents, the agent that requests a service is called a *client* while the agent that receives a service request is called a *server*. In every period, each agent in the community is involved in two matches, one as a client and the other as a server. Note that the agent with whom an agent interacts as a client can be different from that with whom he interacts as a server, reflecting asymmetric interests between a pair of agents in a given instant.

We model the interaction between a pair of matched agents as a gift-giving game [32]. In a gift-giving game, the server has the binary choice of whether to fulfill or decline the request, while the client has no choice. The server's action determines the payoffs of both agents. If the server fulfills the client's request, the client receives a service benefit of $b > 0$ while the server suffers a service cost of $c > 0$. We assume that $b > c$ so that the service of an agent creates a positive net social benefit. If the server declines the request, both agents receive zero payoffs. The set of actions for the server is denoted by $\mathcal{A} = \{F, D\}$, where *F* stands for "fulfill" and *D* for "decline". The payoff matrix of the gift-giving game is presented in Table 1. An agent plays the gift-giving game repeatedly with changing partners until he leaves the community. We assume that at the end of each period a fraction $\alpha \in [0,1]$ of agents in the current population leave and the same amount of new agents join the community. We refer to α as the *turnover rate* [17].

Social welfare in a time period is measured by the average payoff of the agents in that period. Since $b > c$, social welfare is maximized when all the servers choose action *F* in the gift-giving games they play, which yields payoff $b - c$ to every agent. On the contrary, action *D* is the dominant strategy for the server in the gift-giving game, which constitutes a Nash equilibrium of the gift-giving game. When every server chooses his action to maximize his current payoff myopically, an inefficient outcome arises where every agent receives zero payoff.

TABLE 2. Payoff matrix of a gift-giving game.

	Server: *F*	Server: *D*
Client	b, $-c$	0, 0

B. Incentive Schemes Based on a Rating Protocol

In order to improve the efficiency of the myopic equilibrium, we use incentive schemes based on *rating*

[6] It should be noted that our analysis can be readily extended to the case where each agent generates a service request with a probability $\lambda < 1$. We assume $\lambda = 1$ in this paper only for the simplicity of illustrations.

[7] Uniform random matching serves as a good assumption and is commonly adopted in the analysis for online communities with large populations where agents provide identical services, e.g. wireless relay networks [2], peer-to-peer networks [3][6]. The impact of matching schemes on the incentive of agents and the performance of online communities falls out the scope of this paper, but serves as an important next step in this line of research.

protocols. A rating protocol is defined as the rules that a community uses to regulate the behavior of its members. These rules indicate the established and approved ways of "operating" (e.g., exchanging services) in the community: adherence to these rules is positively rewarded, while failure to follow these rules results in (possibly severe) punishments [36]. This gives rating protocols a potential to provide incentives for cooperation. We consider a rating protocol that consists of a *rating scheme* and a *recommended strategy*, as in [26] and [27]. A rating scheme determines the ratings of agents depending on their past actions as a server, while a recommended strategy prescribes the actions that servers should take depending on the ratings of the matched agents.

Formally, a rating scheme is represented by three parameters (Θ, K, τ): Θ denotes the set of rating scores that an agent can hold, $K \in \Theta$ denotes the initial rating score attached to newly joining agents, and τ is the rating update rule. After a server takes an action, the client sends a report (or feedback) about the action of the server to the third-party device or infrastructure that manages the rating scores of agents, but the report is subject to errors with a small probability ε. That is, with probability ε, *D* is reported when the server takes action *F*, and vice versa. Assuming a binary set of reports, it is without loss of generality to restrict ε in $[0, 1/2]$. When $\varepsilon = 1/2$, reports are completely random and do not contain any meaningful information about the actions of servers. We consider a rating scheme that updates the rating score of a server based only on the rating scores of matched agents and the reported action of the server. Then, a rating scheme can be represented by a mapping $\tau : \Theta \times \Theta \times \mathcal{A} \to \Theta$, where $\tau(\theta, \tilde{\theta}, a_R)$ is the new rating score for a server with current rating score θ when he is matched with a client with rating score $\tilde{\theta}$ and his action is reported as a_R. A recommended strategy is represented by a mapping $\sigma : \Theta \times \Theta \to \mathcal{A}$, where $\sigma(\theta, \tilde{\theta})$ is the approved action for a server with rating score θ that is matched with a client with rating score $\tilde{\theta}$. [8]

To simplify our analysis, we initially impose the following restrictions on rating schemes. [9]

1) Θ is a nonempty finite set, i.e., $\Theta = \{0, 1, \dots, L\}$ for some nonnegative integer L.
2) $K = L$.
3) τ is defined by

$$\tau(\theta, \tilde{\theta}, a_R) = \begin{cases} \min\{\theta + 1, L\} & \text{if } a_R = \sigma(\theta, \tilde{\theta}), \\ 0 & \text{if } a_R \neq \sigma(\theta, \tilde{\theta}). \end{cases} \tag{1}$$

Note that with the above three restrictions a nonnegative integer L completely describes a rating scheme, and thus a rating protocol can be represented by a pair $\kappa = (L, \sigma)$. We call the rating scheme determined by L the *maximal punishment rating scheme* (MPRS) with punishment length L. In the MPRS with punishment length L, there are $L+1$ rating scores, and the initial rating score is specified as L. If the reported action of the server is the same as that specified by the recommended strategy σ, the server's rating score is increased by 1 while not exceeding L. Otherwise, the server's rating score is set as 0. A schematic representation of an MPRS is provided

[8] The strategies in the existing rating mechanisms [26][27] determine the server's action based solely on the client's rating score, and thus can be considered as a special case of the recommended strategies proposed in this paper.

[9] We will relax the second and third restrictions in Section V.

in Fig. 1.

Below we summarize the sequence of events in a time period:

1) Agents generate service requests and are matched.
2) Each server observes the rating of his client and then determines his action.
3) Each client reports the action of his server.
4) The rating scores of agents are updated, and each agent observes his new rating score for the next period.
5) A fraction of agents leave the community, and the same amount of new agents join the community.

III. PROBLEM FORMULATION

A. Stationary Distribution of Rating Scores

As time passes, the rating scores of agents are updated and agents leave and join the community. Thus, the distribution of rating scores in the community evolves over time. Let $\eta^t(\theta)$ be the fraction of θ-agents in the total population at the beginning of an arbitrary period t, where a θ-agent means an agent with rating θ. Suppose that all the agents in the community follow a given recommended strategy σ. Then the transition from $\{\eta^t(\theta)\}_{\theta=0}^{L}$ to $\{\eta^{t+1}(\theta)\}_{\theta=0}^{L}$ is determined by the rating scheme, taking into account the turnover rate α and the error probability ε, as shown in the following expressions:

$$\begin{aligned}
&\eta^{t+1}(0) = (1-\alpha)\varepsilon, \\
&\eta^{t+1}(\theta) = (1-\alpha)(1-\varepsilon)\eta^t(\theta-1) \quad \text{for } 1 \le \theta \le L-1, \\
&\eta^{t+1}(L) = (1-\alpha)(1-\varepsilon)\{\eta^t(L)+\eta^t(L-1)\}+\alpha.
\end{aligned} \tag{2}$$

Since we are interested in the long-term payoffs of the agents, we study the distribution of rating scores in the long run.

Definition 1 (Stationary distribution) $\{\eta(\theta)\}$ is a stationary distribution of rating scores under the dynamics defined by (2) if it satisfies $\sum_{\theta=0}^{L}\eta(\theta)=1$, $\eta(\theta)\ge 0, \forall\theta$, and

$$\begin{aligned}
&\eta(0) = (1-\alpha)\varepsilon, \\
&\eta(\theta) = (1-\alpha)(1-\varepsilon)\eta(\theta-1) \quad \text{for } 1 \le \theta \le L-1, \\
&\eta(L) = (1-\alpha)(1-\varepsilon)\{\eta(L)+\eta(L-1)\}+\alpha.
\end{aligned} \tag{3}$$

The following lemma shows the existence of and convergence to a unique stationary distribution.

Lemma 1. For any $\varepsilon \in [0,1/2]$ and $\alpha \in [0,1]$, there exists a unique stationary distribution $\{\eta(\theta)\}$ whose expression is given by

$$\begin{aligned}
&\eta(\theta) = (1-\alpha)^{\theta+1}(1-\varepsilon)^{\theta}\varepsilon, \ \textit{for} \ 0 \le \theta \le L-1, \\
&\eta(L) = \begin{cases} 1 & \textit{if } \alpha=\varepsilon=0, \\ \dfrac{(1-\alpha)^{L+1}(1-\varepsilon)^{L}\varepsilon+\alpha}{1-(1-\alpha)(1-\varepsilon)} & \textit{otherwise.} \end{cases}
\end{aligned} \tag{4}$$

Moreover, the stationary distribution $\{\eta(\theta)\}$ is reached within $(L+1)$ periods starting from any initial distribution.

Proof: Suppose that $\alpha > 0$ or $\varepsilon > 0$. Then (3) has a unique solution

$$\eta(\theta) = (1-\alpha)^{\theta+1}(1-\varepsilon)^{\theta}\varepsilon,\ \textit{for}\ 0 \le \theta \le L-1, \\ \eta(L) = \frac{(1-\alpha)^{L+1}(1-\varepsilon)^{L}\varepsilon + \alpha}{1-(1-\alpha)(1-\varepsilon)}, \tag{5}$$

which satisfies $\sum_{\theta=0}^{L}\eta(\theta) = 1$. Suppose that $\alpha = 0$ and $\varepsilon = 0$. Then solving (3) together with $\sum_{\theta=0}^{L}\eta(\theta) = 1$ yields a unique solution $\eta\left(\theta\right) = 0$ for $0 \le \theta \le L-1$ and $\eta(L) = 1$. It is easy to see from the expressions in (2) that $\eta(\theta)$ is reached within $(\theta+1)$ periods, for all θ, starting from any initial distribution. ■

Since the coefficients in the equations that define a stationary distribution are independent of the recommended strategy that the agents follow, the stationary distribution is also independent of the recommended strategy, as can be seen in (4). Thus, we will write the stationary distribution as $\{\eta_L(\theta)\}$ to emphasize its dependence on the rating scheme, which is represented by L.

B. Sustainable Rating Protocols

We now investigate the incentive of agents to follow a prescribed recommended strategy. For simplicity, we check the incentive of agents at the stationary distribution of rating scores, as in [27] and [31]. Since we consider a non-cooperative scenario, we need to check whether an agent can improve his long-term payoff by a unilateral deviation. Note that any unilateral deviation from an individual agent would not affect the evolution of rating scores and thus the stationary distribution, [10] because we consider a continuum of agents.

Let $c_\sigma(\theta,\tilde{\theta})$ be the cost suffered by a server with rating score θ that is matched with a client with rating score $\tilde{\theta}$ and follows a recommended strategy σ, i.e., $c_\sigma(\theta,\tilde{\theta}) = c$ if $\sigma(\theta,\tilde{\theta}) = F$ and $c_\sigma(\theta,\tilde{\theta}) = 0$ if $\sigma(\theta,\tilde{\theta}) = D$. Similarly, let $b_\sigma(\theta,\tilde{\theta})$ be the benefit received by a client with rating score $\tilde{\theta}$ that is matched with a server with rating score θ following a recommended strategy σ, i.e., $b_\sigma(\theta,\tilde{\theta}) = b$ if $\sigma(\theta,\tilde{\theta}) = F$ and $b_\sigma(\theta,\tilde{\theta}) = 0$ if $\sigma(\theta,\tilde{\theta}) = D$. Since we consider uniform random matching, the expected period payoff of a θ-agent under rating protocol κ before he is matched is given by

$$v_\kappa(\theta) = \sum_{\tilde{\theta}\in\Theta}\eta_L(\tilde{\theta})b_\sigma(\tilde{\theta},\theta) - \sum_{\tilde{\theta}\in\Theta}\eta_L(\tilde{\theta})c_\sigma(\tilde{\theta},\theta)\,. \tag{6}$$

To evaluate the long-term payoff of an agent, we use the discounted sum criterion in which the long-term payoff of an agent is given by the expected value of the sum of discounted period payoffs from the current period. Let $p_\kappa(\theta' \mid \theta)$ be the transition probability that a θ-agent becomes a θ'-agent in the next period under rating protocol κ. Since we consider MPRS, $p_\kappa(\theta' \mid \theta)$ can be expressed as

[10] This is true for any deviation by agents of measure zero.

$$p_{\kappa}(\theta' \mid \theta) = \begin{cases} 1-\varepsilon & \text{if } \theta' = \min\{\theta+1, L\}, \\ \varepsilon & \text{if } \theta' = 0, \\ 0 & \text{otherwise}, \end{cases} \quad \text{for all } \theta \in \Theta . \tag{7}$$

Then we can compute the long-term payoff of an agent from the current period (before he is matched) by solving the following recursive equations

$$v_{\kappa}^{\infty}(\theta) = v_{\kappa}(\theta) + \delta \sum_{\theta' \in \Theta} p_{\kappa}(\theta' \mid \theta) v_{\kappa}^{\infty}(\theta') \quad \text{for all } \theta \in \Theta , \tag{8}$$

where $\delta = \beta(1-\alpha)$ is the weight that an agent puts on his future payoff. Since an agent leaves the community with probability α at the end of the current period, the expected future payoff of a θ-agent is given by $(1-\alpha)\sum_{\theta' \in \Theta} p_{\kappa}(\theta' \mid \theta) v_{\kappa}^{\infty}(\theta')$, assuming that an agent receives zero payoff once he leaves the community. The expected future payoff is multiplied by a common discount factor $\beta \in [0,1)$, which reflects the time preference, or patience, of agents.

Now suppose that an agent deviates and uses a strategy σ' under rating protocol κ. Since the deviation of a single agent does not affect the stationary distribution, the expected period payoff of a deviating θ-agent is given by

$$v_{\kappa,\sigma'}(\theta) = \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) b_{\sigma}(\tilde{\theta}, \theta) + \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) c_{\sigma'}(\theta, \tilde{\theta}) . \tag{9}$$

Let $p_{\kappa,\sigma'}(\theta' \mid \theta, \tilde{\theta})$ be the transition probability that a θ-agent using the strategy σ' becomes a θ'-agent in the next period under rating protocol κ, when he is matched with a client with rating score $\tilde{\theta}$. For each θ, $\theta' = \min\{\theta+1, L\}$ with probability $(1-\varepsilon)$ and $\theta' = 0$ with probability ε if $\sigma(\theta, \tilde{\theta}) = \sigma'(\theta, \tilde{\theta})$ while the probabilities are reversed otherwise. Then $p_{\kappa,\sigma'}(\theta' \mid \theta) = \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) p_{\kappa,\sigma'}(\theta' \mid \theta, \tilde{\theta})$ gives the transition probability of a θ-agent before knowing the rating score of his client, and the long-term payoff of a deviating agent from the current period (before he is matched) can be computed by solving

$$v_{\kappa,\sigma'}^{\infty}(\theta) = v_{\kappa,\sigma'}(\theta) + \delta \sum_{\theta' \in \Theta} p_{\kappa,\sigma'}(\theta' \mid \theta) v_{\kappa,\sigma'}^{\infty}(\theta') \quad \text{for all } \theta \in \Theta . \tag{10}$$

In our model, a server decides whether to provide a service or not after he is matched with a client and observes the rating score of the client. Hence, we check the incentive for a server to follow a recommended strategy at the point when he knows the rating score of the client. Suppose that a server with rating score θ is matched with a client with rating score $\tilde{\theta}$. When the server follows the recommended strategy σ prescribed by rating protocol κ, he receives the long-term payoff $-c_{\sigma}(\theta, \tilde{\theta}) + \delta \sum_{\theta'} p_{\kappa}(\theta' \mid \theta) v_{\kappa}^{\infty}(\theta')$, excluding the possible benefit as a client in the current period. On the contrary, when the server deviates to a recommended strategy σ', he receives the long-term payoff $-c_{\sigma'}(\theta, \tilde{\theta}) + \delta \sum_{\theta'} p_{\kappa,\sigma'}(\theta' \mid \theta, \tilde{\theta}) v_{\kappa,\sigma'}^{\infty}(\theta')$, again excluding the possible benefit as a client. By comparing these two payoffs, we can check whether a θ-agent has an incentive to deviate to σ' when he is matched with a client with rating score $\tilde{\theta}$.

Definition 2 (Sustainable rating protocols) A rating protocol κ is *sustainable* if

$$-c_\sigma(\theta,\tilde{\theta}) + \delta \sum\nolimits_{\theta'} p_\kappa(\theta' \mid \theta) v_\kappa^\infty(\theta') \geq -c_{\sigma'}(\theta,\tilde{\theta}) + \delta \sum\nolimits_{\theta'} p_{\kappa,\sigma'}(\theta' \mid \theta,\tilde{\theta}) v_{\kappa,\sigma'}^\infty(\theta') \quad (11)$$

for all σ', for all $(\theta,\tilde{\theta})$.

In words, a rating protocol $\kappa = (L,\sigma)$ is sustainable if no agent can gain from a unilateral deviation regardless of the rating score of the client he is matched with when every other agent follows recommended strategy σ and the rating scores are determined by the MPRS with punishment length *L*. Thus, under a sustainable rating protocol, agents follow the prescribed recommended strategy in their self-interest. Checking whether a rating protocol is sustainable using the above definition requires computing deviation gains from all possible recommended strategies, whose computation complexity can be quite high for moderate values of *L*. By employing the criterion of unimprovability in Markov decision theory [38], we establish the one-shot deviation principle for sustainable rating protocols, which provides simpler conditions. For notation, let c_a be the cost suffered by a server that takes action a, and let $p_{\kappa,a}(\theta' \mid \theta,\tilde{\theta})$ be the transition probability that a θ-agent becomes a θ'-agent in the next period under rating protocol κ when he takes action a to a client with rating score $\tilde{\theta}$. The values of $p_{\kappa,a}(\theta' \mid \theta,\tilde{\theta})$ can be obtained in a similar way to obtain $p_{\kappa,\sigma'}(\theta' \mid \theta,\tilde{\theta})$, by comparing a with $\sigma(\theta,\tilde{\theta})$.

Lemma 2 (One-shot Deviation Principle). A rating protocol κ is sustainable if and only if

$$c_\sigma(\theta,\tilde{\theta}) - c_a \leq \delta \left[\sum_{\theta'} \{ p_\kappa(\theta' \mid \theta) - p_{\kappa,a}(\theta' \mid \theta,\tilde{\theta}) \} v_\kappa^\infty(\theta') \right] \quad (12)$$

for all $a \neq \sigma(\theta,\tilde{\theta})$, for all $(\theta,\tilde{\theta})$.

Proof: If rating protocol κ is sustainable, then clearly there are no profitable one-shot deviations. We can prove the converse by showing that, if κ is not sustainable, there is at least one profitable one-shot deviation. Since $c_\sigma(\theta,\tilde{\theta})$ and c_a are bounded, this is true by the unimprovability property in Markov decision theory [33][34]. ■

Lemma 2 shows that if an agent cannot gain by unilaterally deviating from σ only in the current period and following σ afterwards, he cannot gain by switching to any other recommended strategy σ' either, and vice versa. The left-hand side of (12) can be interpreted as the current gain from choosing a, while the right-hand side of (12) represents the discounted expected future loss due to the different transition probabilities induced by choosing a. Using the one-shot deviation principle, we can derive incentive constraints that characterize sustainable rating protocols.

First, consider a pair of rating scores $(\theta,\tilde{\theta})$ such that $\sigma(\theta,\tilde{\theta}) = F$. If the server with rating score θ serves the client, he suffers the service cost of c in the current period while his rating score in the next period becomes $\min\{\theta+1, L\}$ with probability $(1-\varepsilon)$ and 0 with probability ε. Thus, the expected long-term payoff of a θ-

agent when he provides a service is given by

$$V_\theta(F;F) = -c + \delta[(1-\varepsilon)v_\kappa^\infty(\min\{\theta+1,L\}) + \varepsilon v_\kappa^\infty(0)] \tag{13}$$

On the contrary, if a θ-agent deviates and declines the service request, he avoids the cost of c in the current period while his rating score in the next period becomes 0 with probability $(1-\varepsilon)$ and $\min\{\theta+1,L\}$ with probability ε. Thus, the expected long-term payoff of a θ-agent when he does not provide a service is given by

$$V_\theta(D;F) = \delta[(1-\varepsilon)v_\kappa^\infty(0) + \varepsilon v_\kappa^\infty(\min\{\theta+1,L\})]. \tag{14}$$

The incentive constraint that a θ-agent does not gain from a one-shot deviation is given by $V_\theta(F;F) \geq V_\theta(D;F)$, which can be expressed as

$$\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \geq c. \tag{15}$$

Now, consider a pair of rating scores $(\theta,\tilde{\theta})$ such that $\sigma(\theta,\tilde{\theta}) = D$. Using a similar argument as above, we can show that the incentive constraint that a θ-agent does not gain from a one-shot deviation can be expressed as

$$\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \geq -c. \tag{16}$$

Note that (15) implies (16), and thus for θ such that $\sigma(\theta,\tilde{\theta}) = F$ for some $\tilde{\theta}$, we can check only the first incentive constraint (15). Therefore, a rating protocol κ is sustainable if and only if (15) holds for all θ such that $\sigma(\theta,\tilde{\theta}) = F$ for some $\tilde{\theta}$ and (16) holds for all θ such that $\sigma(\theta,\tilde{\theta}) = D$ for all $\tilde{\theta}$. The left-hand side of the incentive constraints (15) and (16) can be interpreted as the loss from punishment that rating protocol κ applies to a θ-agent for not following the recommended strategy. Therefore, in order to induce a θ-agent to provide a service to some clients, the left-hand side should be at least as large as the service cost c, which can be interpreted as the deviation gain. We use $\min_{\theta\in\Theta}\{\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)]\}$ to measure the strength of the *incentive for cooperation* under rating protocol κ, where cooperation means providing the requested service in our context.

C. Rating Protocol Design Problem

Since we assume that the community operates at the stationary distribution of rating scores, social welfare under rating protocol κ can be computed by

$$U_\kappa = \sum_\theta \eta_L(\theta) v_\kappa(\theta). \tag{17}$$

The community operator aims to choose a rating protocol that maximizes social welfare among sustainable rating protocols. Then the problem of designing a rating protocol can be formally expressed as

$$\begin{aligned} &\underset{(L,\sigma)}{\text{maximize}} \quad U_\kappa = \sum_\theta \eta_L(\theta) v_\kappa(\theta) \\ &\text{subject to} \quad \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \geq c,\ \forall\theta \ \textit{such that}\ \exists\tilde{\theta}\ \textit{such that}\ \sigma(\theta,\tilde{\theta}) = F, \\ &\qquad\qquad \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \geq -c,\ \forall\theta\ \textit{such that}\ \sigma(\theta,\tilde{\theta}) = D\ \forall\tilde{\theta}. \end{aligned} \tag{18}$$

A rating protocol that solves the design problem (18) is called an *optimal rating protocol*.

IV. ANALYSIS OF OPTIMAL RATING PROTOCOLS

A. Optimal Value of the Design Problem

We first investigate whether there exists a sustainable rating protocol, i.e., whether the design problem (18) has a feasible solution. Fix the punishment length L and consider a recommended strategy σ_L^D defined by $\sigma_L^D(\theta,\tilde{\theta}) = D$ for all $(\theta,\tilde{\theta})$. Since there is no service provided in the community when all the agents follow σ_L^D, we have $v_{(L,\sigma_L^D)}^{\infty}(\theta) = 0$ for all θ. Hence, the relevant incentive constraint (16) is satisfied for all θ, and the rating protocol (L,σ_L^D) is sustainable. This shows that the design problem (18) always has a feasible solution.

Assuming that an optimal rating protocol exists, let U^* be the optimal value of the design problem (18). In the following proposition, we study the properties of U^*.

Proposition 1. The optimal value of the design problem (18) satisfies the following properties:

(i) $0 \le U^* \le b - \dfrac{1-\varepsilon}{1-2\varepsilon}c$.

(ii) $U^* = 0$ if $\dfrac{c}{b} > \dfrac{\beta(1-\alpha)(1-2\varepsilon)}{1-\beta(1-\alpha)(2-3\varepsilon)}$.

(iii) $U^* \ge [1-(1-\alpha)\varepsilon](b-c)$ if $\dfrac{c}{b} \le \beta(1-\alpha)(1-2\varepsilon)$.

(iv) $U^* = b - c$ if $\varepsilon = 0$ and $\dfrac{c}{b} \le \beta(1-\alpha)$.

(v) $U^* = b - c$ only if $\varepsilon = 0$ and $\dfrac{c}{b} \le \dfrac{\beta(1-\alpha)}{1-\beta(1-\alpha)}$.

Proof: See Appendix A. ■

Proposition 1(i) proves that the optimal social welfare cannot be negative but is always strictly bounded away from $b-c$, which is the social welfare when all agents cooperates, when $\varepsilon > 0$. Hence full cooperation cannot be achieved in this scenario. Since we obtain zero social welfare at myopic equilibrium, without using a rating protocol, we are interested in whether we can sustain a rating protocol in which agents cooperate in a positive proportion of matches. In other words, we look for conditions on the parameters $(b,c,\beta,\alpha,\varepsilon)$ that yield $U^* > 0$. From Proposition 1(ii) and (iii), we can regard $c/b \le [\beta(1-\alpha)(1-2\varepsilon)]/[1-\beta(1-\alpha)(2-3\varepsilon)]$ and $c/b \le \beta(1-\alpha)(1-2\varepsilon)$ as necessary and sufficient conditions for $U^* > 0$, respectively. Moreover, when there are no report errors (i.e., $\varepsilon = 0$), we can interpret $c/b \le \beta(1-\alpha)/[1-\beta(1-\alpha)]$ and $c/b \le \beta(1-\alpha)$ as necessary and sufficient conditions to achieve the maximum social welfare $U^* = b-c$, respectively. As a corollary of Proposition 1, we obtain the following results in the limit.

Corollary 1. For any (b,c) such that $b > c$, (i) U^* converges to $b-c$ as $\beta \to 1$, $\alpha \to 0$, and $\varepsilon \to 0$, and (ii) U^* converges to 0 as $\beta \to 0$, $\alpha \to 1$, or $\varepsilon \to 1/2$. ■

Corollary 1 shows that we can design a sustainable rating protocol that achieves near efficiency (i.e., U^* close to $b-c$) when the community conditions are good (i.e., β is close to 1, and α and ε are close to 0). Moreover, it suffices to use only two ratings (i.e., $L=1$) for the design of such a rating protocol. On the contrary, no cooperation can be sustained (i.e., $U^*=0$) when the community conditions are bad (i.e., β is close to 0, α is close to 1, or ε is close to 1/2), as implied by Proposition 1(ii).

B. Optimal Recommended strategies Given a Punishment Length

In order to obtain analytical results, we consider the design problem (18) with a fixed punishment length L, denoted DP_L. Note that DP_L has a feasible solution, namely σ_L^D, for any L and that there are a finite number (total $2^{(L+1)^2}$) of possible recommended strategies given L. Therefore, DP_L has an optimal solution for any L. We use U_L^* and σ_L^* to denote the optimal value and the optimal recommended strategy of DP_L, respectively. We first show that increasing the punishment length cannot decrease the optimal value.

Proposition 2. $U_L^* \geq U_{L'}^*$ for all L and L' such that $L \geq L'$.

Proof: See Appendix B. ■

Proposition 2 shows that U_L^* is non-decreasing in L. Since $U_L^* < b-c$ when $\varepsilon > 0$, we have $U^* = \lim_{L\to\infty} U_L^* = \sup_L U_L^*$. It may be the case that the incentive constraints eventually prevent the optimal value from increasing with L so that the supremum is attained by some finite L. This conjecture is verified in Fig. 2, where U_L^* stops increasing when $L \geq 5$. Hence, it is plausible for the protocol designer to set an upper bound on L in practical designs with little efficiency loss incurred. Now we analyze the structure of optimal recommended strategies given a punishment length. The properties characterized in the following proposition can effectively reduce the design space of the optimal recommended strategy given L and thus reduces the computation complexity of the optimal rating protocol design.

Proposition 3. If we have that $\varepsilon > 0$ and $\alpha < 1$, the optimal rating protocol exhibits the following structures:

(i) A 0-agent does not receive service from some agents, i.e., $\sigma_L^*(\theta, 0) = D,\ \exists \theta \in \Theta$.

(ii) If $\sigma_L^*(0,\hat{\theta}) = F$ for some $\hat{\theta}$, then agents with sufficiently high rating scores always receive service from 0-agents, i.e., $\sigma_L^*(0,\tilde{\theta}) = F$ for all $\tilde{\theta} \geq \min\{\ln\frac{c}{b} / \ln\beta, L\}$.

(iii) *L*- agents receive service from other agents whose rating scores are sufficiently high, i.e., if $\theta \in \{1,\ldots,L-1\}$ satisfies $\theta \geq L - (\ln\frac{c}{b} - \ln Y(\alpha,\varepsilon,L)) \big/ \ln\beta$, where

$$Y(\alpha,\varepsilon,L) = \frac{(1-\alpha)^{L+1}(1-\varepsilon)^L\varepsilon - (1-\alpha)^{L+2}(1-\varepsilon)^{L+1}\varepsilon}{(1-\alpha)^{L+1}(1-\varepsilon)^L\varepsilon + \alpha}, \tag{19}$$

then $\sigma_L^*(\theta, L) = F$.

(iv) L-agents always provide service to other L-agents, i.e., $\sigma_L^*(L, L) = F$.

Proof: See Appendix C. ■

As Proposition 3 shows, to construct an optimal rating protocol, sufficient punishment should be provided to agents with low rating scores while sufficient rewards should be provided to agents with high rating scores.

C. Illustration with L=1 and L=2

We can represent a recommended strategy σ_L as an $(L+1)\times(L+1)$ matrix whose (i,j)-entry is given by $\sigma_L(i-1, j-1)$. Proposition 3 provides some structures of an optimal recommended strategy σ_L^* in the first column and the last row of the matrix representation, but it does not fully characterize the solution of DP_L. Here we aim to obtain the solution of DP_L for $L=1$ and 2 and analyze how it changes with the parameters. We first begin with the case of two ratings, i.e., $L=1$. In this case, if $\sigma_1(\theta,\tilde{\theta}) = F$ for some $(\theta,\tilde{\theta})$, the relevant incentive constraint to sustain $\kappa = (1, \sigma_1)$ is $\delta(1-2\varepsilon)[v_\kappa^\infty(1) - v_\kappa^\infty(0)] \geq c$. By Proposition 3(ii) and (iv), if $\sigma_1^*(\theta,\tilde{\theta}) = F$ for some $(\theta,\tilde{\theta})$, then $\sigma_1^*(0,1) = \sigma_1^*(1,1) = F$, provided that $\varepsilon > 0$ and $\alpha < 1$. Hence, among the total of 16 possible recommended strategies, only four can be optimal recommended strategies. These four recommended strategies are

$$\sigma_1^1 = \begin{bmatrix} D & F \\ F & F \end{bmatrix},\ \sigma_1^2 = \begin{bmatrix} F & F \\ D & F \end{bmatrix},\ \sigma_1^3 = \begin{bmatrix} D & F \\ D & F \end{bmatrix},\ \sigma_1^4 = \sigma_1^D = \begin{bmatrix} D & D \\ D & D \end{bmatrix}. \tag{20}$$

For notational convenience, we define a recommended strategy σ_L^{D0} by $\sigma_L^{D0}(\theta, 0) = D$ for all θ and $\sigma_L^{D0}(\theta,\tilde{\theta}) = F$ for all θ and all $\tilde{\theta} > 0$. In (20), we have $\sigma_1^3 = \sigma_1^{D0}$. The following proposition specifies the optimal recommended strategy given the parameters.

Proposition 4. Suppose that $0 < (1-\alpha)\varepsilon < 1/2$. Then

$$\sigma_1^* = \begin{cases} \sigma_1^1 & \text{if } 0 < \dfrac{c}{b} \leq \dfrac{\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon}{1+\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon}, \\ \sigma_1^2 & \text{if } \dfrac{\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon}{1+\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon} < \dfrac{c}{b} \leq \dfrac{\beta(1-\alpha)(1-2\varepsilon)[1-(1-\alpha)\varepsilon]}{1-\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon}, \\ \sigma_1^3 & \text{if } \dfrac{\beta(1-\alpha)(1-2\varepsilon)[1-(1-\alpha)\varepsilon]}{1-\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon} < \dfrac{c}{b} \leq \beta(1-\alpha)(1-2\varepsilon), \\ \sigma_1^4 & \text{if } \beta(1-\alpha)(1-2\varepsilon) < \dfrac{c}{b} < 1. \end{cases} \tag{21}$$

Proof: Let $\kappa^i = (1, \sigma_1^i)$, for $i = 1,2,3,4$. We obtain that

$$\begin{aligned} U_{\kappa^1} &= (1-\eta_1(0)^2)(b-c), \quad U_{\kappa^2} = (1-\eta_1(0)\eta_1(1))(b-c), \\ U_{\kappa^3} &= (1-\eta_1(0))(b-c), \quad U_{\kappa^4} = 0. \end{aligned} \tag{22}$$

Since $0 < (1-\alpha)\varepsilon < 1/2$, we have $\eta_1(0) < \eta_1(1)$. Thus, we have $U_{\kappa^1} > U_{\kappa^2} > U_{\kappa^3} > U_{\kappa^4}$. Also, we obtain that

$$\begin{aligned} &v_{\kappa^1}^{\infty}(1) - v_{\kappa^1}^{\infty}(0) = \eta_1(0)(b-c), \quad && v_{\kappa^2}^{\infty}(1) - v_{\kappa^2}^{\infty}(0) = b - \eta_1(0)(b-c), \\ &v_{\kappa^3}^{\infty}(1) - v_{\kappa^3}^{\infty}(0) = b, && v_{\kappa^4}^{\infty}(1) - v_{\kappa^4}^{\infty}(0) = 0. \end{aligned} \tag{23}$$

Thus, we have $v_{\kappa^3}^{\infty}(1) - v_{\kappa^3}^{\infty}(0) > v_{\kappa^2}^{\infty}(1) - v_{\kappa^2}^{\infty}(0) > v_{\kappa^1}^{\infty}(1) - v_{\kappa^1}^{\infty}(0) > v_{\kappa^4}^{\infty}(1) - v_{\kappa^4}^{\infty}(0)$. By choosing the recommended strategy that yields the highest social welfare among feasible ones, we obtain the result. ■

Proposition 4 shows that the optimal recommended strategy is determined by the service cost-to-benefit ratio, c/b. When c/b is sufficiently small, the recommended strategy σ_1^1 can be sustained, yielding the highest social welfare among the four candidate recommended strategies. As c/b increases, the optimal recommended strategy changes from σ_1^1 to σ_1^2 to σ_1^3 and eventually to σ_1^4. Fig. 3 shows the optimal recommended strategies with $L=1$ as c varies. The parameters we use to obtain the results in the figures of this paper are set as follows unless otherwise stated: $\beta = 0.8$, $\alpha = 0.1$, $\varepsilon = 0.2$, and $b = 10$. Fig. 3(a) plots the incentive for cooperation of the four recommended strategies. We can find the region of c in which each strategy is sustained by comparing the incentive for cooperation with the service cost c for σ_1^1, σ_1^2, and σ_1^3, and with $-c$ for σ_1^4. The solid portion of the lines indicates that the strategy is sustained while the dashed portion indicates that the strategy is not sustained. Fig. 3(b) plots the social welfare of the four candidate strategies, with solid and dashed portions having the same meanings. The triangle-marked line represents the optimal value, which takes the maximum of the social welfare of all sustained strategies.

Next, we analyze the case of three ratings, i.e., $L=2$. In order to provide a partial characterization of the optimal recommended strategy σ_2^*, we introduce the following notation. Let $\sigma_2^{\#}$ be the recommended strategy with $L=2$ that maximizes $\min\{v_{\kappa}^{\infty}(1) - v_{\kappa}^{\infty}(0), v_{\kappa}^{\infty}(2) - v_{\kappa}^{\infty}(0)\}$ among all the recommended strategies with $L=2$. Let $\gamma \triangleq \delta(1-\varepsilon)$ as defined in Appendix A, and define a recommended strategy σ_L^B by $\sigma_L^B(L-1, 0) = D$ and $\sigma_L^B(\theta, \tilde{\theta}) = F$ for all $(\theta, \tilde{\theta}) \neq (L-1, 0)$. We have the following conclusion about σ_2^* and $\sigma_2^{\#}$.

Proposition 5. Suppose that $\varepsilon > 0$, $\alpha < 1$, and

$$\frac{c}{b} < \frac{\eta_2(2)}{\eta_2(1)} \frac{1-\gamma}{\gamma} < \frac{b}{c}. \tag{24}$$

(i) $\sigma_2^{\#} = \sigma_2^{D0}$; (ii) if $\eta_2(0) < \eta_2(2)$, then $\sigma_2^* = \sigma_2^B$.

Proof: (i) Let $\kappa = (2, \sigma_2^{D0})$. Then $v_{\kappa}^{\infty}(1) - v_{\kappa}^{\infty}(0) = v_{\kappa}^{\infty}(2) - v_{\kappa}^{\infty}(0) = b$. We can show that, under the given conditions, any change from σ_2^{D0} results in a decrease in the value of $v_{\kappa}^{\infty}(1) - v_{\kappa}^{\infty}(0)$, which proves that σ_2^{D0} maximizes $\min\{v_{\kappa}^{\infty}(1) - v_{\kappa}^{\infty}(0), v_{\kappa}^{\infty}(2) - v_{\kappa}^{\infty}(0)\}$.

(ii) Since $\varepsilon > 0$ and $\alpha < 1$, we have $\eta_2(\theta) > 0$ for all $\theta = 0,1,2$, and thus replacing *D* with *F* in an element of a recommended strategy always improves social welfare. Hence, we first consider the recommended strategy σ_L^F defined by $\sigma_L^F(\theta,\tilde{\theta}) = F$ for all $(\theta,\tilde{\theta})$. σ_2^F maximizes social welfare U_κ among all the recommended strategies with $L = 2$, but $v_\kappa^\infty(1) - v_\kappa^\infty(0) = v_\kappa^\infty(2) - v_\kappa^\infty(0) = 0$. Thus, we cannot find parameters such that σ_2^F satisfies the incentive constraints, and thus $\sigma_2^F \neq \sigma_2^*$. Now consider recommended strategies in which there is exactly one *D* element. We can show that, under the given conditions, having $\sigma_2(\theta,\tilde{\theta}) = D$ at $(\theta,\tilde{\theta})$ such that $\tilde{\theta} > 0$ yields $v_\kappa^\infty(1) - v_\kappa^\infty(0) < 0$, whereas having $\sigma_2(\theta,\tilde{\theta}) = D$ at $(\theta,\tilde{\theta})$ such that $\tilde{\theta} = 0$ yields both $v_\kappa^\infty(1) - v_\kappa^\infty(0) > 0$ and $v_\kappa^\infty(2) - v_\kappa^\infty(0) > 0$. Thus, for any recommended strategy having the only *D* element at $(\theta,\tilde{\theta})$ such that $\tilde{\theta} > 0$, there do not exist parameters in the considered parameter space with which the incentive constraint for 0-agents, $\delta(1-2\varepsilon)\left[v_\kappa^\infty(1) - v_\kappa^\infty(0)\right] \geq c$, is satisfied. On the other hand, for any recommended strategy having the only *D* element at $(\theta,\tilde{\theta})$ such that $\tilde{\theta} = 0$, we can satisfy both incentive constraints by choosing $\beta > 0$, $\alpha < 1$, $\varepsilon < 1/2$, and c sufficiently close to 0. This shows that, among the recommended strategies having exactly one *D* element, only those having *D* in the first column are possibly sustainable. Since $\eta_2(1) < \eta_2(0) < \eta_2(2)$, σ_2^B achieves the highest social welfare among the three candidate recommended strategies. ■

Let us try to better understand now what the conditions in Proposition 5 mean. Proposition 5(i) implies that the maximum incentive for cooperation that can be achieved with three ratings is $\beta(1-\alpha)(1-2\varepsilon)b$. Hence, cooperation can be sustained with $L = 2$ if and only if $\beta(1-\alpha)(1-2\varepsilon)b \geq c$. That is, if $c/b > \beta(1-\alpha)(1-2\varepsilon)$, then σ_2^D is the only feasible recommended strategy and thus $U_2^* = 0$. Therefore, when we increase c while holding other parameters fixed, we can expect that σ_2^* changes from σ_2^{D0} to σ_2^D around $c = \beta(1-\alpha)(1-2\varepsilon)b$. Note that the same is observed with $L = 1$ in Proposition 4. We can see that $[\eta_\tau(2)/\eta_\tau(1)][(1-\gamma)/\gamma]$ converges to 1 as α goes to 0 and β goes to 1. Hence, for given values of b, c, and ε, the condition (24) is satisfied and thus some cooperation can be sustained if α and β are sufficiently close to 0 and 1, respectively.

Consider a rating protocol $\kappa = (2, \sigma_2^B)$. We obtain that

$$\min\{v_\kappa^\infty(1) - v_\kappa^\infty(0), v_\kappa^\infty(2) - v_\kappa^\infty(0)\} = v_\kappa^\infty(2) - v_\kappa^\infty(0) = (1-\alpha)^2(1-\varepsilon)\varepsilon(b - \beta c) \tag{25}$$

and $U_\kappa = (1-(1-\alpha)^3(1-\varepsilon)\varepsilon^2)(b-c)$. Proposition 5(ii) is stating that $\sigma_2^* = \sigma_2^B$ when the community conditions are "favorable." More precisely, we have $\sigma_2^* = \sigma_2^B$ if $(1-\alpha)^2(1-\varepsilon)\varepsilon(b-\beta c) \geq c$, or

$$\frac{c}{b} \leq \frac{\beta(1-\alpha)^3(1-2\varepsilon)(1-\varepsilon)\varepsilon}{1+\beta^2(1-\alpha)^3(1-2\varepsilon)(1-\varepsilon)\varepsilon}. \tag{26}$$

Also, Proposition 5(ii) implies that $U_2^* \leq (1-(1-\alpha)^3(1-\varepsilon)\varepsilon^2)(b-c)$ always holds.

In Fig. 4, we show the optimal value and the optimal recommended strategy of *DP₂* as we vary c. The optimal recommended strategy σ_2^* changes in the following order before becoming σ_2^D as c increases:

$$\sigma_2^1 = \begin{bmatrix} F & F & F \\ D & F & F \\ F & F & F \end{bmatrix}, \sigma_2^2 = \begin{bmatrix} D & F & F \\ F & F & F \\ F & F & F \end{bmatrix}, \sigma_2^3 = \begin{bmatrix} D & F & F \\ D & F & F \\ F & F & F \end{bmatrix},$$
$$\sigma_2^4 = \begin{bmatrix} F & F & F \\ F & F & F \\ D & F & F \end{bmatrix}, \sigma_2^5 = \begin{bmatrix} F & F & F \\ D & F & F \\ D & F & F \end{bmatrix}, \sigma_2^6 = \begin{bmatrix} D & F & F \\ F & F & F \\ D & F & F \end{bmatrix}, \sigma_2^7 = \begin{bmatrix} D & F & F \\ D & F & F \\ D & F & F \end{bmatrix}. \quad (27)$$

Note that $\sigma_2^1 = \sigma_2^B$ for small c and $\sigma_2^7 = \sigma_2^{D0}$ for large c (but not too large to sustain cooperation), which are consistent with the discussion about Proposition 5. For the intermediate values of c, only the elements in the first column change in order to increase the incentive for cooperation. We find that the order of the optimal recommended strategies between $\sigma_2^1 = \sigma_2^B$ and $\sigma_2^7 = \sigma_2^{D0}$ depends on the community's parameters $(b, c, \beta, \alpha, \varepsilon)$.

V. EXTENSIONS

A. Rating Schemes with Shorter Punishment Length

So far we have focused on MPRS under which any deviation in reported actions results in the rating score of 0. Although this class of rating schemes is simple in that a rating scheme can be identified with the number of rating scores, it may not yield the highest social welfare among all possible rating schemes when there are report errors. When there is no report error, i.e., $\varepsilon = 0$, an agent maintains rating score L as long as he follows the prescribed recommended strategy. Thus, in this case, punishment exists only as a threat and it does not result in an efficiency loss. On the contrary, when $\varepsilon > 0$ and $\alpha < 1$, there exist a positive proportion of agents with ratings 0 to $L-1$ in the stationary distribution even if all the agents follow the recommended strategy. Thus, there is a tension between efficiency and incentive. In order to sustain a rating protocol, we need to provide a strong punishment so that agents do not gain by deviation. At the same time, too severe a punishment reduces social welfare. This observation suggests that, in the presence of report errors, it is optimal to provide incentives just enough to prevent deviations. If we can provide a weaker punishment while sustaining the same recommended strategy, it will improve social welfare. One way to provide a weaker punishment is to use a random punishment. For example, we can consider a rating scheme under which the rating score of a θ-agent becomes 0 in the next period with probability $q_\theta \in (0,1]$ and remains the same with probability $1-q_\theta$ when he reportedly deviates from the recommended strategy. By varying the punishment probability q_θ for θ-agents, we can adjust the severity of the punishment applied to θ-agents. This class of rating schemes can be identified by $(L, \{q_\theta\})$. MPRS can be considered as a special case where $q_\theta = 1$ for all θ.

Another way to provide a weaker punishment is to use a smaller punishment length, denoted M. Under the rating scheme with $(L+1)$ rating scores and punishment length M, rating scores are updated by

$$\tau(\theta,\tilde{\theta},a_R)=\begin{cases}\min\{\theta+1,L\} & if\ \ a_R=\sigma(\theta,\tilde{\theta}),\\ \max\{\theta-M,0\} & if\ \ a_R\neq\sigma(\theta,\tilde{\theta}).\end{cases} \tag{28}$$

When a θ-agent reportedly deviates from the recommended strategy, his rating score is reduced by M in the next period if $\theta\geq M$ and becomes 0 otherwise. Note that this rating scheme is analogous to real-world rating schemes for credit rating and auto insurance risk rating. This class of rating schemes can be identified by (L,M) with $1\leq M\leq L$.[11] MPRS can be considered as a special case where $M=L$.

In this paper, we focus on the second approach to investigate the impacts of the punishment length on the social welfare U_κ and the incentive for cooperation $\min_\theta\{\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\})-v_\kappa^\infty(\max\{\theta-M,0\})]\}$ of a rating protocol κ, which is now defined as (L,M,σ). The punishment length *M* affects the evolution of the rating distribution, and the stationary distribution of rating scores with the rating scheme (L,M), $\{\eta_{(L,M)}(\theta)\}_{\theta=0}^L$, satisfies the following equations:

$$\begin{aligned}
&\eta_{(L,M)}(0)=(1-\alpha)\varepsilon\sum_{\theta=0}^{M}\eta_{(L,M)}(\theta),\\
&\eta_{(L,M)}(\theta)=(1-\alpha)(1-\varepsilon)\eta_{(L,M)}(\theta-1)+(1-\alpha)\varepsilon\eta_{(L,M)}(\theta+M)\quad for\ 1\leq\theta\leq L-M,\\
&\eta_{(L,M)}(\theta)=(1-\alpha)(1-\varepsilon)\eta_{(L,M)}(\theta-1)\quad for\ L-M+1\leq\theta\leq L-1,\\
&\eta_{(L,M)}(L)=(1-\alpha)(1-\varepsilon)\{\eta_{(L,M)}(L)+\eta_{(L,M)}(L-1)\}+\alpha.
\end{aligned} \tag{29}$$

Let $\{\mu_{(L,M)}(\theta)\}_{\theta=1}^L$ be the cumulative distribution of $\{\eta_{(L,M)}(\theta)\}_{\theta=1}^L$, i.e., $\mu_{(L,M)}(\theta)=\sum_{k=0}^{\theta}\eta_{(L,M)}(k)$ for $\theta=0,\ldots,L$. Fig. 5 plots the stationary distribution $\{\eta_{(L,M)}(\theta)\}_{\theta=1}^L$ and its cumulative distribution $\{\mu_{(L,M)}(\theta)\}_{\theta=1}^L$ for $L=5$ and $M=1,\ldots,5$. We can see that the cumulative distribution monotonically decreases with M, i.e., $\mu_{(L,M_1)}(\theta)\leq\mu_{(L,M_2)}(\theta)$ for all θ if $M_1>M_2$. This shows that, as the punishment length increases, there are more agents holding a lower rating score. As a result, when the community adopts a recommended strategy that treats an agent with a higher rating score better, increasing the punishment length reduces social welfare while it increases the incentive for cooperation. This trade-off is illustrated in Fig. 6, which plots social welfare and the incentive for cooperation under a rating protocol $(3,M,\sigma_3^C)$ for $M=1,2,3$, where the recommended strategy σ_L^C is defined by $\sigma_L^C(\theta,\tilde{\theta})=F$ if and only if $\tilde{\theta}\geq\theta$, for all θ.

In general, the recommended strategy adopted in the community is determined together with the rating scheme in order to maximize social welfare while satisfying the incentive constraints. The design problem with variable

[11] We can further generalize this class by having the punishment length depend on the rating. That is, when a θ-agent reportedly deviates from the recommended strategy, his rating is reduced to $\theta-M_\theta$ in the next period for some $M_\theta\leq\theta$.

punishment lengths can be formulated as follows. First, note that the expected period payoff of a θ-agent, $v_\kappa(\theta)$, can be computed by (6), with the modification of the stationary distribution to $\{\eta_{(L,M)}(\theta)\}_{\theta=1}^{L}$. Agents' long-term payoffs can be obtained by solving (8), with the transition probabilities now given by

$$p_\kappa(\theta' \mid \theta) = \begin{cases} 1-\varepsilon & if\ \theta' = \min\{\theta+1, L\}, \\ \varepsilon & if\ \theta' = \max\{\theta - M, 0\},\ for\ all\ \theta \in \Theta. \\ 0 & otherwise, \end{cases} \tag{30}$$

Finally, the design problem can be written as

$$\begin{aligned} &\underset{(L,M,\sigma)}{\max imize} \quad U_\kappa = \sum_\theta \eta_{(L,M)}(\theta) v_\kappa(\theta) \\ &subject\ to \quad \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(\max\{\theta-M, 0\})] \geq c, \\ &\qquad\qquad \forall\theta\ such\ that\ \exists\tilde{\theta}\ such\ that\ \sigma(\theta,\tilde{\theta}) = F, \\ &\qquad\qquad \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(\max\{\theta-M, 0\})] \geq -c, \\ &\qquad\qquad \forall\theta\ such\ that\ \sigma(\theta,\tilde{\theta}) = D\ \forall\tilde{\theta}. \end{aligned} \tag{31}$$

We find the optimal recommended strategy given a rating scheme (L, M) for $L = 3$ and $M = 1,2,3$, and plot the social welfare and the incentive for cooperation of the optimal recommended strategies in Fig. 7. Since different values of M induce different optimal recommended strategies given the value of L, there are no monotonic relationships between the punishment length and social welfare as well as the incentive for cooperation, unlike in Fig. 6. The optimal punishment length given L can be obtained by taking the punishment length that yields the highest social welfare, which is plotted in Fig. 8. We can see that, as the service cost c increases, the optimal punishment length increases from 1 to 2 to 3 before cooperation becomes no longer sustainable. This result is intuitive in that larger c requires a stronger incentive for cooperation, which can be achieved by having a larger punishment length.

B. Whitewash-Proof Rating protocols

So far we have restricted our attention to rating schemes where newly joining agents are endowed with the highest rating score, i.e., $K = L$, without worrying about the possibility of whitewashing. We now make the initial rating score *K* as a choice variable of the design problem while assuming that agents can whitewash their rating scores in order to obtain rating score *K* [17]. At the end of each period, agents can decide whether to whitewash their rating scores or not after observing their rating scores for the next period. If an agent chooses to whitewash his rating score, then he leaves and re-joins the community with α fraction of agents and receives initial rating score K. The cost of whitewashing is denoted by $c_w \geq 0$.

The incentive constraints in the design problem (18) are aimed at preventing agents from deviating from the prescribed recommended strategy. In the presence of potential whitewashing attempts, we need additional incentive constraints to prevent agents from whitewashing their rating scores. A rating protocol κ is *whitewash-proof* if and only if $v_\kappa^\infty(K) - v_\kappa^\infty(\theta) \le c_w$ for all $\theta = 0,\ldots,L$.[12] Note that $v_\kappa^\infty(K) - v_\kappa^\infty(\theta)$ is the gain from whitewashing for an agent whose rating score is updated as θ. If $v_\kappa^\infty(K) - v_\kappa^\infty(\theta) \le c_w$, there is no net gain from whitewashing for a θ-agent. We measure the *incentive for whitewashing* under a rating protocol κ by $\max_{\theta\in\Theta}\{v_\kappa^\infty(K) - v_\kappa^\infty(\theta)\}$. A rating protocol is more effective in preventing whitewashing as the incentive for whitewashing is smaller.

To simplify our analysis, we fix the punishment length at $M = L$ so that a rating scheme is represented by (L,K) with $0 \le K \le L$. Let $\{\eta_{(L,K)}(\theta)\}_{\theta=1}^{L}$ be the stationary distribution of rating scores under rating scheme (L,K). Then the design problem is modified as follows (it should be noted here that similar to Section V.A, both $\{v_\kappa(\theta)\}$ and $\{v_\kappa^\infty(\theta)\}$ in this section are computed using a stationary distribution different than (4), which depends on the value of K)

$$\begin{aligned}
&\underset{(L,K,\sigma)}{\text{maximize}} \quad U_\kappa = \sum_\theta \eta_{(L,K)}(\theta) v_\kappa(\theta) \\
&\text{subject to} \quad \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \ge c, \ \forall\theta \text{ such that } \exists\tilde\theta \text{ such that } \sigma(\theta,\tilde\theta) = F, \\
&\qquad\qquad \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \ge -c, \ \forall\theta \text{ such that } \sigma(\theta,\tilde\theta) = D \ \forall\tilde\theta, \\
&\qquad\qquad v_\kappa^\infty(K) - v_\kappa^\infty(\theta) \le c_w, \ \forall\theta.
\end{aligned} \tag{32}$$

Now an optimal rating protocol is the one that maximizes social welfare among sustainable and whitewash-proof rating protocols. Note that the design problem (32) always has a feasible solution for any $c_w \ge 0$ since (L,K,σ_L^D) is sustainable and whitewash-proof for all (L,K). Nevertheless, (L,K,σ_L^D) is trivial since no service takes place in the community as a consequence. Next, we show that given the existence of sustainable rating protocols which deliver a positive level of cooperation, i.e., when U^* solved by (18) is positive, whitewash-proof rating protocols also exist.

Lemma 3. If a rating protocol (L,K,σ) is sustainable, then the rating protocol $(L,0,\sigma)$ is also sustainable. ■

Lemma 3 shows that it never reduces agents' incentive of cooperation by assigning the newly joined agents the lowest ratings. With this result, we prove the existence of whitewash-proof rating protocols

Proposition 6. If $U^* > 0$, then whitewash-proof rating protocols always exist.

Proof: If $U^* > 0$, then sustainable rating protocols that stimulate positive levels of cooperation always exist. According to Lemma 3, if a protocol κ with $K = \theta > 0$ is sustainable, then a protocol κ' with the same rating scheme and recommended strategy and $K = 0$ is also sustainable. Meanwhile, it can be verified that κ' is whitewash-proof since an agent cannot get any benefit by leaving and rejoining the community while also suffering the whitewashing cost. Hence, Proposition 6 follows. ■

[12] This condition assumes that an agent can whitewash his rating only once in his lifespan in the community. More generally, we can consider the case where an agent can whitewash his rating multiple times. For example, an agent can use a deterministic stationary decision rule for whitewashing, which can be represented by a function $w : \Theta \to \{0,1\}$, where $w(\theta) = 1$ (resp. $w(\theta) = 0$) means that the agent whitewashes (resp. does not whitewash) his rating if he holds rating θ in the next period. This will yield a different expression for the gain from whitewashing.

Now we investigate the impacts of the initial rating score K on social welfare and the incentive for whitewashing. We first consider the case where the recommended strategy is fixed. Fig. 9 plots social welfare and the incentive for whitewashing under a rating protocol $(3, K, \sigma_3^C)$ for $K = 0, \ldots, 3$, where σ_L^C is defined by $\sigma_L^C(\theta, \tilde{\theta}) = F$ if and only if $\tilde{\theta} \geq \theta$, for all θ, as in the above section. We can see that larger *K* yields higher social welfare and at the same time a larger incentive for whitewashing since new agents are treated better. Hence, there is a trade-off between efficiency and whitewash-proofness as we increase *K* while fixing the recommended strategy. Next we consider the optimal recommended strategy given a rating scheme (L, K). Fig. 10 plots social welfare and the incentive for whitewashing under the optimal recommended strategy for $L = 3$ and $K = 0, \ldots, 3$. We can see that giving the highest rating score to new agents ($K = 3$) yields the highest social welfare but it can prevent whitewashing only for small values of c. With our parameter specification, choosing $K = 3$ is optimal only for small c, and optimal *K* drops to 0 for other values of c with which some cooperation can be sustained. Fig. 11 plots the optimal initial rating K^* as we vary the whitewashing cost c_w, for $c = 1, 2, 3$. As c_w increases, the incentive constraints for whitewashing becomes less binding, and thus K^* is non-decreasing in c_w. On the other hand, as c increases, it becomes more difficult to sustain cooperation while the difference between $v_\kappa^\infty(0)$ and $v_\kappa^\infty(\min\{\theta + 1, L\})$ increases for all θ such that $\sigma(\theta, \tilde{\theta}) = F$ for some $\tilde{\theta}$. As a result, K^* is non-increasing in c.

C. One-sided Rating Protocol

The above discussion focuses on the design of optimal rating protocols where the recommended strategy utilizes both the rating scores of the client and the server in order to determine the server's action. We refer to such recommended strategies as *two-sided recommended strategies* since they involve the rating scores of both players involved in the stage game.

In this section, we discuss the design of optimal rating protocols with a simple class of recommended strategies that only utilize one-sided rating scores, which we refer to as *one-sided recommended strategies*. Particularly, a one-sided recommended strategy determines an agent's serving action solely based on either the agent's own rating score or the rating score of his client. To differentiate it with the previously discussed two-sided recommended strategies, we denote a one-sided recommended strategy by φ, which can be represented by a mapping $\varphi : \Theta \to \mathcal{A}$, and the corresponding rating protocol, which is called as a one-sided rating protocol, by π. It should be noted that for a one-sided recommended strategy φ that utilizes the clients' rating scores, it is equivalent to a two-sided recommended strategy σ if $\sigma(\theta, \tilde{\theta}) = \varphi(\tilde{\theta})$, $\forall \theta \in \Theta$ and $\forall \tilde{\theta} \in \Theta$. Similarly, for a one-sided recommended strategy φ' that utilizes the servers' rating scores, it is equivalent to a two-sided recommended strategy σ' if $\sigma'(\theta, \tilde{\theta}) = \varphi'(\theta)$, $\forall \theta \in \Theta$ and $\forall \tilde{\theta} \in \Theta$. Therefore, the one-sided recommended strategies represent a subset of the class of two-sided recommended strategies. In this section, we investigate the emerging protocol designs which can be found using such simpler strategies and the corresponding efficiency loss compared to the optimal performance obtained in Section IV.

We first analyze one-side recommended strategies utilizing the clients' rating scores. Given a one-sided recommended strategy φ, the expected period payoff of a θ-agent before he is matched, which is still denoted as $v_{\kappa}(\theta)$ with a little abuse of notation, is given as follows

$$v_{\pi}(\theta) = bI(\varphi(\theta) = F) - \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) c I(\varphi(\tilde{\theta}) = F), \tag{33}$$

where $I(x)$ is an indicator function which takes value of 1 when $x = 1$. The corresponding social welfare, which is denoted as W_{π}, can be computed by

$$W_{\pi} = \sum_{\theta \in \Theta} \eta_L(\theta) v_{\pi}(\theta) = \sum_{\theta \in \Theta} \eta_L(\theta)(b - c) I(\varphi(\theta) = F). \tag{34}$$

The following proposition characterizes the general designing rule of the optimal rating protocol, which is denoted as $\pi^* = (L^*, \varphi^*)$, with the corresponding optimal social welfare denoted as W^*.

Proposition 7. With one-sided recommended strategies utilizing the clients' rating scores, the optimal rating protocol π^* that maximizes (34) satisfies the following conditions: when $c / b \leq \delta(1 - 2\varepsilon)$, $W^* = (1 - \varepsilon + \varepsilon\alpha)(b - c)$ with $L^* = 1$, $\varphi^*(0) = D$, and $\varphi^*(1) = F$; when $c / b > \delta(1 - 2\varepsilon)$, $W^* = 0$ with $\varphi^*(\theta) = D$, $\forall \theta \in \Theta$.

Proof: See Appendix D. ■

The optimal rating protocol from Proposition 7 is surprisingly simple and intuitive. When the cost-to-benefit ratio is sufficiently small such that a positive level of cooperation can be sustained in the community, the optimal rating protocol contains only two different rating values. Modulo the effects of noise, agents who comply with the recommended strategy in the previous period have rating score 1, while agents who deviate from the recommended strategy have rating score 0. The recommended strategy then says that agents should play a tit-for-tat-like strategy, providing services to agents with rating score 1 and punishing those with rating score 0. On the other hand, when the cost-to-benefit ratio is sufficiently large, no cooperation can be sustained and $W^* = 0$. Hence, when the recommended strategy is one-sided and solely utilizes the clients' rating scores, there is no need to construct complicated rating protocols that are difficult for agents to understand, or to heavily optimize parameters of the rating protocol based on the properties of the community. Regarding the fact that the optimal social welfare U^* achieved by two-sided recommended strategies is upper-bounded by $b - c$, we have the following corollary.

Corollary 2. When $c / b \leq \delta(1 - 2\varepsilon)$, $U^* - W^* < \varepsilon(1 - \alpha)(b - c)$. ■

Therefore, the efficiency loss introduced by one-sided recommended strategies monotonically decreases and approaches 0 when $\varepsilon \to 0$ or $\alpha \to 1$.

We then study one-sided recommended strategies utilizing only the servers' rating scores. It is shown in the following proposition that no cooperation can be sustained in this case, which always yields an optimal social welfare $W^* = 0$. Hence, one-sided recommended strategies utilizing only the servers' rating scores can never correctly incentivize service provisions and prevent agents from free-riding.

Proposition 8. With one-sided recommended strategies utilizing the servers' rating scores, the optimal rating protocol π^* always delivers an optimal social welfare $W^* = 0$ with $\varphi^*(\theta) = D$, $\forall \theta \in \Theta$.

Proof: See Appendix D. ■

VI. Illustrative Example

In this section, we present numerical results to illustrate in detail the performance of optimal rating protocols. Unless stated otherwise, the setting of the community is as follows: the benefit per service ($b = 10$), the cost per service ($c = 1$), the discount factor ($\beta = 0.8$), the turnover rate ($\alpha = 0.1$), the report error ($\varepsilon = 0.2$), the punishment step ($M = L$), and the initial rating score ($K = L$). Since the number of all possible recommended strategies given a punishment length L increases exponentially with L, it takes a long time to compute the optimal recommended strategy even for a moderate value of L. Hence, we consider rating protocols $\kappa = (L, \sigma_L^*)$ for $L = 1,2,3$.

We first compare the performances of the optimal rating protocol and the fixed rating protocol for $L = 1,2,3$. For each L, we use (L, σ_L^C) as the fixed rating protocol. Fig. 12 illustrates the results, with the black bar representing the pareto optimal value $b - c$, i.e., the highest social welfare that can be possibly sustained by a rating protocol, the gray bar representing the social welfare of the optimal rating protocol, and the white bar representing the social welfare of (L, σ_L^C). As it shows, the optimal rating protocol (L, σ_L^*) outperforms (L, σ_L^C). When c is small, (L, σ_L^*) delivers higher social welfare than (L, σ_L^C). When c is sufficiently large such that no cooperation can be sustained under (L, σ_L^C) (the height of the white bar becomes 0), a positive level of cooperation can still be sustained under (L, σ_L^*).

Next, we analyze the impacts of the community's parameters on the performance of optimal rating protocols.

Impact of the Discount Factor: We discuss the impact of the discount factor β on the performance of optimal rating protocols. As β increases, an agent puts a higher weight on his future payoff relative to his instant payoff. Hence, with larger β, it is easier to sustain cooperation using future reward and punishment through a rating protocol. This is illustrated in Fig. 13(a), which shows that social welfare is non-decreasing in β.

Impact of the Turnover Rate: Increasing α has two opposite effects on social welfare. As α increases, the weight on the future payoffs, $\delta = \beta(1 - \alpha)$, decreases, and thus it becomes more difficult to sustain cooperation. On the other hand, as α increases, there are more agents holding the highest rating score given the restriction $K = L$. In general, agents with the highest rating score are treated well under optimal recommended strategies, which implies a positive effect of increasing α on social welfare. From Fig. 13(b), we can see that, when α is large, the first effect is dominant, making cooperation not sustainable. We can also see that the second effect is dominant for the values of α with which cooperation can be sustained, yielding an increasing tendency of social welfare with respect to α.

Impact of the Report Errors: As ε increases, it becomes more difficult to sustain cooperation because reward and punishment provided by a rating protocol becomes more random. At the same time, larger ε increases the fraction of 0-agents in the stationary distribution, which usually receive the lowest long-term payoff among all rating scores. Therefore, we can expect that optimal social welfare has a non-increasing tendency with respect to ε, as illustrated in Fig. 13(c). When ε is sufficiently close to 1/2, σ_L^D is the only sustainable recommended strategy and social welfare falls to 0. On the other direction, as ε approaches 0, social welfare converges to its upper bound $b-c$, regardless of the punishment length, as can be seen from Proposition 1(iii). We can also observe from Fig. 13 that the regions of α and ε where some cooperation can be sustained (i.e., $U_L^* > 0$) become wider as L increases, whereas that of β is independent of L.

VII. COMPARISONS WITH EXISTING WORKS ON REPEATED GAMES WITH IMPERFECT MONIROTING

In this section, we compare our work with the existing literature on repeated games with imperfect monitoring, including the works on reputation systems [23]-[25] and the seminal work of Fudenberg, Levine and Maskin [35]. Importantly, our work exhibits significant technical differences from the existing literature. We would like to point out that although the results derived in this work exhibit some structural similarity to [23]-[25], they are derived under completely different settings and using different analytical methods. Also, the methodology in [35] cannot be applied in our work.

We first compare our work with [23]-[25]. The models in [23]-[25] assume that agents have fixed roles in the community (i.e. sellers and buyers), which is common in applications where the groups of sellers and buyers are separated and usually do not overlap. However, in our work, agents are symmetric in the sense that each of them can play both roles of server and of client. Our model is more appropriate to characterize resource/knowledge sharing online communities.

More importantly, the objective in the protocol design of our work is also different from [23]-[25]. The design objective in [23]-[25] is to maximize the expected discounted long-term payoff of the seller starting from the best reputation and a clean history (e.g. Propositions 2 and 3 in [23]). Translating this into the mathematical representation adopted in our paper, the objective function in [23]-[25] is

$$\underset{(L,\sigma)}{\text{maximize}} \quad U_\kappa = v_\kappa^\infty(L) \tag{35}$$

Such objective function is reasonable in the setting of [23]-[25] because they assume that there is a unique long-lived player (seller) in the game who is foresighted. In each period, the seller selects one buyer to interact with. Therefore, it makes sense to focus on the life-time discounted utility of this seller starting from the moment it joins the community. Also, it should be pointed out that the selection of this objective function (35) is the main reason why the optimal design can be achieved in [23]-[25] using a two-level reputation set with $L=1$. Since the future utility is discounted, the optimum of (35) can be achieved using a simple grim-trigger strategy, i.e. the seller cooperates as long as his reputation remains at “good” and does not cooperate when his reputation falls to “bad”.

In contrast, our work assumes multiple long-lived players coexisting in the community (i.e. each agent is long-lived and foresighted). Meanwhile, in each period, there are multiple interactions between different players. Hence, it is more reasonable to maximize the average social welfare of all agents in the long-run as defined in (17), which can be proved to be equivalent to the following objective function:

$$\underset{(L,\sigma)}{\text{maximize}} \quad U_\kappa = \sum_\theta \eta_L(\theta) v_\kappa^\infty(\theta) \tag{36}$$

It is easy to observe that given the grim-trigger strategy designed in [23], the long-run timing-average payoff of the seller is actually 0, even though his expected discounted long-term payoff is maximized. Therefore, the design in [23]-[25] is never optimal from a social welfare perspective, given $\varepsilon > 0$.

Finally, [23]-[25] focus on deriving the rating schemes that can achieve the upper bound of (A1) under the condition when the seller is sufficiently patient with his discount factor δ (i.e. β in our paper) close to 1 or when the payment-to-cost ratio ρ (i.e. b / c in our paper) is sufficiently large. They did not provide much insight on how to derive the optimal rating scheme and what is the optimal expected long-term utility that can be achieved when δ and ρ are small (They simply state that the optimal expected long-term utility is below the upper bound of (A1) in this case). In contrast, our work tries to characterize the optimal rating scheme design for the entire region of the parameters $(\beta, b, c, \varepsilon)$, but not only the scenario when these parameters are ideal.

Next, we compare our work with [35]. The model in [35] does not consider the anonymity and random matching among agents. Also, it assumes that the entire history of public signals (i.e. the outcome of each stage game) is revealed during the repeated game. It should be noted that the proof of Proposition 1 in [23] also relies on this assumption on the information structure in order to obtain the upper bound on the efficiency. Nevertheless, under the rating protocol proposed in our work, each agent only observes a limited set of past L signals from the past periods and hence, the assumption that all past signals are revealed no longer holds here.

Another difference between our work and [35] is that [35] considers the limiting case where the discount factor $\beta \to 1$, while our work obtains the optimal design for the general case of $\beta < 1$.

Finally, the objective in [35] is also to maximize the expected long-term utility of players starting from the beginning of the game (as described in (35)), which is different from our objective function (17).

Regarding all the above reasons, we would like to point out that the methodology in [35] cannot be applied to our work.

VIII. CONCLUSIONS

In this paper, we used the idea of rating protocols to establish a rigorous framework for the design and analysis of a class of incentive schemes to sustain cooperation in online communities. We derived conditions for sustainable rating protocols, under which no agent gains by deviating from the prescribed recommended strategy. We formulated the problem of designing an optimal rating protocol and characterized optimal social welfare and optimal recommended strategies given parameters. As special cases, we analyzed the one-sided rating protocol which only utilizes the rating score of one party in the stage-game. It was shown that when only the clients' rating scores are utilized, the optimal one-sided rating protocol preserves a simple structure with two-level rating score, whereas when only the servers' rating scores are utilized, no sustainable one-sided rating protocol can be designed. We also discussed the impacts of punishment lengths and whitewashing possibility on the design and performance of optimal rating protocols, identifying a trade-off between efficiency and incentives. Lastly, we presented numerical results to illustrate the impacts of the discount factor, the turnover rate, and the probability of report errors on the performance of optimal rating protocols. Our framework provides a foundation for designing incentive schemes which can be deployed in real-world communities populated by anonymous, self-interested individuals.

APPENDIX A

PROOF OF PROPOSITION 1

(i) $U^* \geq 0$ follows by noting that (L, σ_L^D) is feasible. It is shown in [35] that in a repeated game with public signal, the maximum long-run player sequential equilibrium payoff when the signal has full support is the solution of the following linear programming problem:

$$\begin{aligned} \underset{a}{\text{maximize}} \quad & u^\infty \\ \text{subject to} \quad & u^\infty = u - c + \delta(1-\varepsilon)v^\infty(+) + \delta\varepsilon v^\infty(-), \quad \textit{for } a = F \\ & u^\infty \geq u - c + \delta(1-\varepsilon)v^\infty(+) + \delta\varepsilon v^\infty(-), \quad \textit{for } a = D \\ & u^\infty \geq u + \delta\varepsilon v^\infty(+) + \delta(1-\varepsilon)v^\infty(-), \quad \textit{for } a = F \\ & u^\infty = u + \delta\varepsilon v^\infty(+) + \delta(1-\varepsilon)v^\infty(-), \quad \textit{for } a = D \end{aligned} . \tag{37}$$

Here, u is the benefit that the player can receive as a client in a stage game, and $v^\infty(+)$ is the expected long-term utility of this player if his rating score is increased and $v^\infty(-)$ is the expected long-term utility when his rating score is decreased. With simple computation, the maximum solution of (37) can be written as $\frac{1}{1-\delta}(b - \frac{1-\varepsilon}{1-2\varepsilon}c)$. Since no player can achieve a long-run payoff higher than this, it can be concluded that the social welfare is also upper-bounded by $b - \frac{1-\varepsilon}{1-2\varepsilon}c$.

(ii) By (8), we can express $v_\kappa^\infty(1) - v_\kappa^\infty(0)$ as

$$\begin{aligned} & v_\kappa^\infty(1) - v_\kappa^\infty(0) \\ &= v_\kappa(1) + \delta[(1-\varepsilon)v_\kappa^\infty(2) + \varepsilon v_\kappa^\infty(0)] - v_\kappa(0) - \delta[(1-\varepsilon)v_\kappa^\infty(1) + \varepsilon v_\kappa^\infty(0)] \\ &= v_\kappa(1) - v_\kappa(0) + \delta(1-\varepsilon)[v_\kappa^\infty(2) - v_\kappa^\infty(1)]. \end{aligned} \tag{38}$$

Similarly, we have

$$\begin{aligned}
v_\kappa^\infty(2) - v_\kappa^\infty(1) &= v_\kappa(2) - v_\kappa(1) + \delta\left(1-\varepsilon\right)\left[v_\kappa^\infty(3) - v_\kappa^\infty(2)\right], \\
&\vdots \\
v_\kappa^\infty(L-1) - v_\kappa^\infty(L-2) &= v_\kappa(L-1) - v_\kappa(L-2) + \delta\left(1-\varepsilon\right)\left[v_\kappa^\infty(L) - v_\kappa^\infty(L-1)\right], \\
v_\kappa^\infty(L) - v_\kappa^\infty(L-1) &= v_\kappa(L) - v_\kappa(L-1).
\end{aligned} \tag{39}$$

In general, for $\theta = 1, \ldots, L$,

$$v_\kappa^\infty(\theta) - v_\kappa^\infty(\theta-1) = \sum_{l=0}^{L-\theta} \gamma^l [v_\kappa(\theta+l) - v_\kappa(\theta+l-1)], \tag{40}$$

where we define $\gamma = \delta(1-\varepsilon)$. Thus, we obtain

$$\begin{aligned}
&v_\kappa^\infty(\theta) - v_\kappa^\infty(0) \\
&= v_\kappa(\theta) - v_\kappa(0) + \gamma[v_\kappa(\theta+1) - v_\kappa(1)] + \cdots + \gamma^{L-\theta}[v_\kappa(L) - v_\kappa(L-\theta)] \\
&\quad + \gamma^{L-\theta+1}[v_\kappa(L) - v_\kappa(L-\theta+1)] + \cdots + \gamma^{L-1}[v_\kappa(L) - v_\kappa(L-1)] \\
&= \sum_{l=0}^{L-1} \gamma^l [v_\kappa(\min\{\theta+l, L\}) - v_\kappa(l)],
\end{aligned} \tag{41}$$

for $\theta = 1, \ldots, L$.

Since $-c \le v_\kappa(\theta) \le b$ for all θ, we have $v_\kappa(\theta) - v_\kappa(\tilde{\theta}) \le b + c$ for all $(\theta, \tilde{\theta})$. Hence, by (41),

$$v_\kappa^\infty(\theta) - v_\kappa^\infty(0) \le \frac{1-\gamma^L}{1-\gamma}(b+c) \le \frac{b+c}{1-\gamma} \tag{42}$$

for all $\theta = 1, \ldots, L$, for all $\kappa = (L, \sigma)$. Therefore, if $\delta(1-2\varepsilon)[(b+c)/(1-\gamma)] < c$, or equivalently, $c/b > [\beta(1-\alpha)(1-2\varepsilon)]/[1-\beta(1-\alpha)(2-3\varepsilon)]$, then the incentive constraint (15) cannot be satisfied for any θ, for any rating protocol (L, σ). This implies that any recommended strategy σ such that $\sigma(\theta, \tilde{\theta}) = F$ for some $(\theta, \tilde{\theta})$ is not feasible, and thus $U^* = 0$.

(iii) For any L, define a recommended strategy σ_L^{D0} by $\sigma_L^{D0}(\theta, \tilde{\theta}) = D$ for $\tilde{\theta} = 0$ and $\sigma_L^{D0}(\theta, \tilde{\theta}) = F$ for all $\tilde{\theta} > 0$, for all θ. In other words, with σ_L^{D0} each agent declines the service request of 0-agents while providing a service to other agents. Consider a rating protocol $\kappa = (1, \sigma_1^{D0})$. Then $v_\kappa(0) = -\eta_1(1)c$ and $v_\kappa(1) = b - \eta_1(1)c$. Hence, $U_\kappa = [1-(1-\alpha)\varepsilon](b-c)$ and $v_\kappa^\infty(1) - v_\kappa^\infty(0) = b$, and thus the incentive constraint $\delta(1-2\varepsilon)(v_\kappa^\infty(1) - v_\kappa^\infty(0)) \ge c$ is satisfied by the hypothesis $c/b \le \beta(1-\alpha)(1-2\varepsilon)$. Since there exists a feasible solution that achieves $U_\kappa = [1-(1-\alpha)\varepsilon](b-c)$, we have $U^* \ge [1-(1-\alpha)\varepsilon](b-c)$.

(iv) Suppose, on the contrary to the conclusion, that $U^* = b - c$. If $\alpha = 1$, then (15) cannot be satisfied for any θ, for any κ, which implies $U^* = 0$. Hence, it must be the case that $\alpha < 1$. Let $\kappa^* = (L^*, \sigma^*)$ be an optimal rating protocol that achieves $U^* = b - c$. Since $\varepsilon > 0$ and $\alpha < 1$, $\eta_{L^*}(\theta) > 0$ for all θ by (4). Since $U^* = U_{\kappa^*} = (b-c)\sum_{\theta,\tilde{\theta}} \eta_{L^*}(\theta)\eta_{L^*}(\tilde{\theta}) I(\sigma^*(\theta,\tilde{\theta}) = F)$, σ^* should have $\sigma^*(\theta,\tilde{\theta}) = F$ for all $(\theta,\tilde{\theta})$. However, under this recommended strategy, all the agents are treated equally, and thus $v_{\kappa^*}^{\infty}(0) = \cdots = v_{\kappa^*}^{\infty}(L^*)$. Then σ^* cannot satisfy the relevant incentive constraint (15) for all θ since the left-hand side of (15) is zero, which contradicts the optimality of (L^*, σ^*).

(v) The result can be obtained by combining (i) and (iii).

(vi) Suppose that $U^* = b - c$, and let (L, σ) be an optimal rating protocol that achieves $U^* = b - c$. By (iv), we obtain $\varepsilon = 0$. Then by (4), $\eta_L(\theta) = 0$ for all $0 \le \theta \le L-1$ and $\eta_L(L) = 1$. Hence, σ should have $\sigma(L, L) = F$ in order to attain $U^* = b - c$. Since $v_\kappa(L) = b - c$ and $v_\kappa(\theta) \ge -c$ for all $0 \le \theta \le L-1$, we have $v_\kappa^\infty(L) - v_\kappa^\infty(0) \le b / (1-\gamma)$ by (41). If $\delta b / (1-\delta) < c$, then the incentive constraint for L-agents, $\delta[v_\kappa^\infty(L) - v_\kappa^\infty(0)] \ge c$, cannot be satisfied. Therefore, we obtain $c / b \le \delta / (1-\delta)$. ■

APPENDIX B

PROOF OF PROPOSITION 2

Choose an arbitrary L. To prove the result, we will construct a recommended strategy σ_{L+1} using punishment length $L+1$ that is feasible and achieves U_L^*. Define σ_{L+1} by

$$\sigma_{L+1}(\theta,\tilde{\theta}) = \begin{cases} \sigma_L^*(\theta,\tilde{\theta}) & \text{for } \theta \le L \text{ and } \tilde{\theta} \le L, \\ \sigma_L^*(L,\tilde{\theta}) & \text{for } \theta = L+1 \text{ and } \tilde{\theta} \le L, \\ \sigma_L^*(\theta, L) & \text{for } \theta \le L \text{ and } \tilde{\theta} = L+1, \\ \sigma_L^*(L, L) & \text{for } \theta = L+1 \text{ and } \tilde{\theta} = L+1. \end{cases} \tag{43}$$

Let $\kappa = (L, \sigma_L^*)$ and $\kappa' = (L+1, \sigma_{L+1})$. From (4), we have $\eta_{L+1}(\theta) = \eta_L(\theta)$ for $\theta = 0, \ldots, L-1$ and $\eta_{L+1}(L) + \eta_{L+1}(L+1) = \eta_L(L)$. Using this and (6), it is straightforward to see that $v_{\kappa'}(\theta) = v_\kappa(\theta)$ for all $\theta = 0, \ldots, L$ and $v_{\kappa'}(L+1) = v_\kappa(L)$. Hence, we have that

$$\begin{aligned} U_{\kappa'} &= \sum_{\theta=0}^{L+1} \eta_{L+1}(\theta) v_{\kappa'}(\theta) = \sum_{\theta=0}^{L-1} \eta_{L+1}(\theta) v_{\kappa'}(\theta) + \sum_{\theta=L}^{L+1} \eta_{L+1}(\theta) v_{\kappa'}(\theta) \\ &= \sum_{\theta=0}^{L-1} \eta_L(\theta) v_\kappa(\theta) + \sum_{\theta=L}^{L+1} \eta_{L+1}(\theta) v_\kappa(L) \\ &= \sum_{\theta=0}^{L-1} \eta_L(\theta) v_\kappa(\theta) + \eta_L(L) v_\kappa(L) = U_\kappa = U_L^*. \end{aligned} \tag{44}$$

Using (41), we can show that $v_{\kappa'}^{\infty}(\theta) - v_{\kappa'}^{\infty}(0) = v_{\kappa}^{\infty}(\theta) - v_{\kappa}^{\infty}(0)$ for all $\theta = 1,\ldots,L$ and $v_{\kappa'}^{\infty}(L+1) - v_{\kappa'}^{\infty}(0) = v_{\kappa}^{\infty}(L) - v_{\kappa}^{\infty}(0)$. By the definition of σ_{L+1}, the right-hand side of the relevant incentive constraint (i.e., c or $-c$) for each $\theta = 0,\ldots,L$ is the same both under σ_L^* and under σ_{L+1}. Also, under σ_{L+1}, the right-hand side of the relevant incentive constraint for $\theta = L+1$ is the same as that for $\theta = L$. Therefore, σ_{L+1} satisfies the relevant incentive constraints for all $\theta = 0,\ldots,L+1$. ■

APPENDIX C

PROOF OF PROPOSITION 3

To facilitate the proof, we define $u_{\kappa}^{\infty}(\theta)$ by

$$u_{\kappa}^{\infty}(\theta) = \sum_{l=0}^{\infty} \gamma^l v_{\kappa}(\min\{\theta + l, L\}) \tag{45}$$

for $\theta = 0,\ldots,L$. Then, by (41), we have $v_{\kappa}^{\infty}(\theta) - v_{\kappa}^{\infty}(0) = u_{\kappa}^{\infty}(\theta) - u_{\kappa}^{\infty}(0)$ for all $\theta = 1,\ldots,L$. Thus, we can use $u_{\kappa}^{\infty}(\theta) - u_{\kappa}^{\infty}(0)$ instead of $v_{\kappa}^{\infty}(\theta) - v_{\kappa}^{\infty}(0)$ in the incentive constraints of DP_L.

Suppose that $\sigma_L^*(0,\hat{\theta}) = F$ for some $\hat{\theta}$. Then the relevant incentive constraint for a 0-agent is $\delta(1-2\varepsilon)[u_{\kappa}^{\infty}(1) - u_{\kappa}^{\infty}(0)] \geq c$. Suppose that $\sigma_L^*(0,\overline{\theta}) = D$ for some $\overline{\theta} \in \{1,\ldots,L-1\}$ such that $\overline{\theta} \geq \ln\frac{c}{b} / \ln\beta$. Consider a recommended strategy σ_L' defined by

$$\sigma_L'(\theta,\tilde{\theta}) = \begin{cases} \sigma_L^*(\theta,\tilde{\theta}) & \text{for } (\theta,\tilde{\theta}) \neq (0,\overline{\theta}), \\ F & \text{for } (\theta,\tilde{\theta}) = (0,\overline{\theta}). \end{cases} \tag{46}$$

That is, σ_L' is the recommended strategy that differs from σ_L^* only at $(0,\overline{\theta})$. Let $\kappa = (L,\sigma_L^*)$ and $\kappa' = (L,\sigma_L')$. Note that $v_{\kappa'}(0) - v_{\kappa}(0) = -\eta_{\tau}(\overline{\theta})c < 0$ and $v_{\kappa'}(\overline{\theta}) - v_{\kappa}(\overline{\theta}) = \eta_{\tau}(0)b > 0$ since $\varepsilon > 0$ and $\alpha < 1$. Thus, $U_{\kappa'} - U_{\kappa} = \eta_L(0)\eta_L(\overline{\theta})(b-c) > 0$. Also,

$$u_{\kappa'}^{\infty}(\theta) - u_{\kappa}^{\infty}(\theta) = \begin{cases} [v_{\kappa'}(0) - v_{\kappa}(0)] + \gamma^{\overline{\theta}}[v_{\kappa'}(\overline{\theta}) - v_{\kappa}(\overline{\theta})] \\ \quad = (1-\alpha)^{\overline{\theta}+1}(1-\varepsilon)^{\overline{\theta}}\varepsilon[\beta^{\overline{\theta}}b - c] & \text{for } \theta = 0, \\ \gamma^{\overline{\theta}-\theta}[v_{\kappa'}(\overline{\theta}) - v_{\kappa}(\overline{\theta})] & \text{for } \theta = 1,\ldots,\overline{\theta}, \\ 0 & \text{for } \theta = \overline{\theta}+1,\ldots,L. \end{cases} \tag{47}$$

Since $\overline{\theta} \geq \ln\frac{c}{b} / \ln\beta$, we have $u_{\kappa'}^{\infty}(0) - u_{\kappa}^{\infty}(0) \leq 0$. Thus, $u_{\kappa'}^{\infty}(\theta) - u_{\kappa'}^{\infty}(0) \geq u_{\kappa}^{\infty}(\theta) - u_{\kappa}^{\infty}(0)$ for all

$\theta = 1,\ldots,L$. Since $\sigma_L^*(0,\hat{\theta}) = F$ for some $\hat{\theta}$, the relevant incentive constraint for a θ-agent is the same both under σ_L^* and under σ_L', for all θ. Hence, σ_L' satisfies the incentive constraints of *DP*$_L$, which contradicts the optimality of σ_L^*. This proves that $\sigma_L^*(0,\tilde{\theta}) = F$ for all $\tilde{\theta} \geq \ln\frac{c}{b} / \ln\beta$. Similar approaches can be used to prove $\sigma_L^*(0,L) = F$, (i), and (iii). ■ φ^*

APPENDIX D

PROOF OF PROPOSITION 7

To prove this proposition, we first show that the recommended strategy in the optimal rating protocol is always threshold based. That is, there is an integer h such that $\varphi^*(\theta) = D$ for all $\theta < h$ and $\varphi^*(\theta) = F$ for all $\theta \geq h$. We use a contradiction to verify this. Suppose $\varphi^*(\tilde{\theta}) = F$ and $\varphi^*(\tilde{\theta}+1) = D$ for some θ. Consider a strategy φ' satisfying $\varphi'(\theta) = \varphi^*(\theta)$ for all $\theta \neq \tilde{\theta}+1$ and $\varphi'(\tilde{\theta}) = \varphi^*(\tilde{\theta})$. Since is sustainable, then according to (8), it is easy to verify that φ' is also sustainable. Meanwhile, φ' delivers a higher social welfare than φ^*. Hence, the fact that φ^* is optimal is contradicted and we can conclude that the optimal recommended strategy is always threshold-based. According to (34), it is obvious that the social welfare upon agents' compliance monotonically decreases with the threshold h. Next, we analyse how h affects the sustainability of the recommended strategy.

Under a threshold-based recommended strategy with threshold h, the expected long-term utility can be recursively represented as

$$\begin{aligned} v_\kappa^\infty(\theta) &= b - \sum_{\theta \geq h} \eta_L(\theta)c + \delta(1-\varepsilon)v_\kappa^\infty(\min\{\theta+1, L\}) + \delta\varepsilon v_\kappa^\infty(\max\{\theta-1, 0\}),\ \ if\ \ \theta \geq h \\ v_\kappa^\infty(\theta) &= -\sum_{\theta \geq h} \eta_L(\theta)c + \delta(1-\varepsilon)v_\kappa^\infty(\min\{\theta+1, L\}) + \delta\varepsilon v_\kappa^\infty(\max\{\theta-1, 0\}),\ \ if\ \ \theta < h \end{aligned} \quad . \quad (48)$$

Hence, it can be shown that $v_\kappa^\infty(1) - v_\kappa^\infty(0) = \min_\theta\{v_\kappa^\infty(\theta) - v_\kappa^\infty(0)\}$. Meanwhile, $v_\kappa^\infty(1) - v_\kappa^\infty(0)$ monotonically decreases against h. Therefore, if a recommended strategy with threshold h is sustainable, then the recommended strategy with threshold $h-1$ is also sustainable. However, according to Proposition 3, it should be noted that the recommended strategy with $h=0$ can never be sustainable. Therefore, to sum up, for a given punishment length *L*, when sustainable one-sided rating protocols exist, the threshold of the optimal recommended strategy φ^* is always $h=1$. It is easy to compute that any threshold-based recommended strategy φ with $h=1$ is sustainable if and only if $c/b \leq \delta(1-2\varepsilon)$ with the resulting social welfare to be $W = (1-\varepsilon+\varepsilon\alpha)(b-c)$. Hence, Proposition 7 follows. ■

PROOF OF PROPOSITION 8

The proof of this proposition is similar to that of Proposition 7. First, it can be proved that in the optimal rating protocol that uses servers' rating scores, the recommended strategy is always threshold-based. That is, there is an integer h such that $\varphi^*(\theta) = F$ for all $\theta < h$ and $\varphi^*(\theta) = D$ for all $\theta \geq h$. However, under such threshold-based recommended strategy, $v_\kappa^\infty(\theta) - v_\kappa^\infty(0) \leq c$ hold for all $\theta \in \Theta$ and hence, (12) is never satisfied and Proposition 6 follows. ■

APPENDIX E

FIGURES

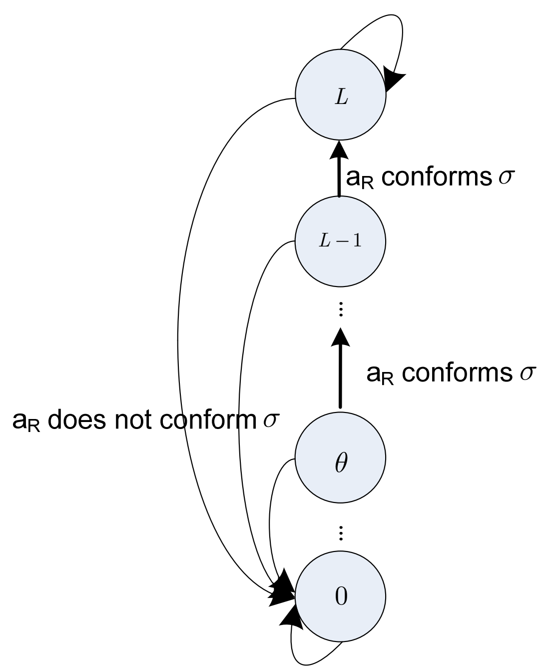


Fig. 1. Schematic representation of a maximal punishment rating scheme.

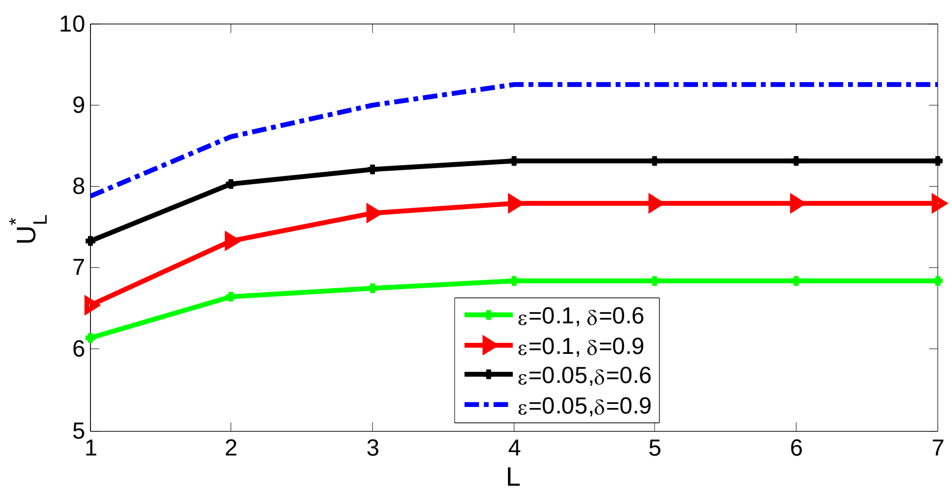
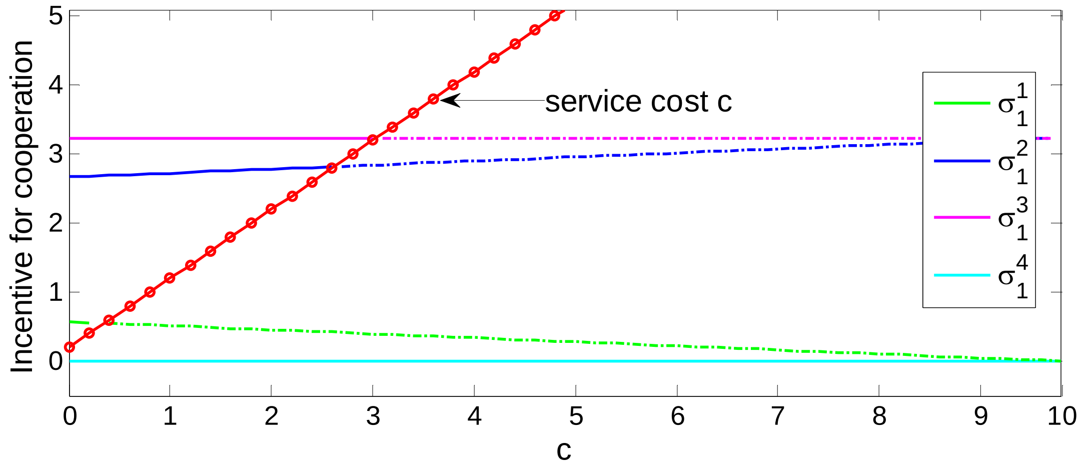
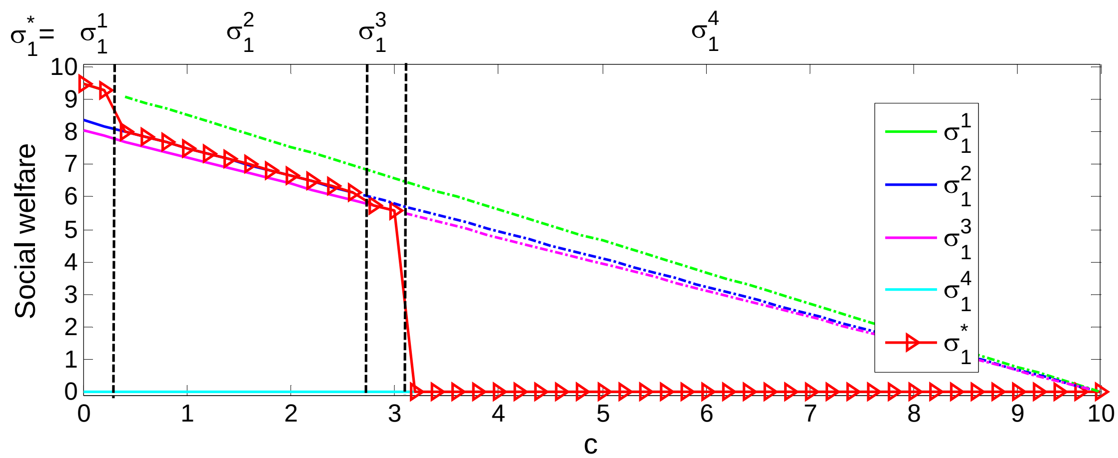


Fig. 2. Optimal performance giving the punishment length *L*.



(a)



(b)

Fig. 3. Performance of the four candidate recommended strategies when $L = 1$: (a) incentive for cooperation, and (b) social welfare and the optimal recommended strategy.

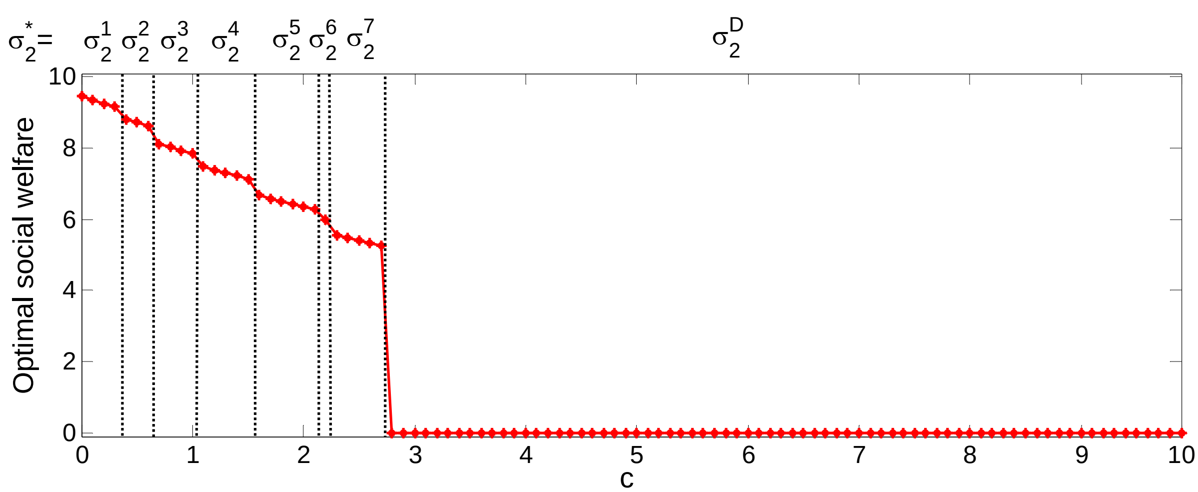
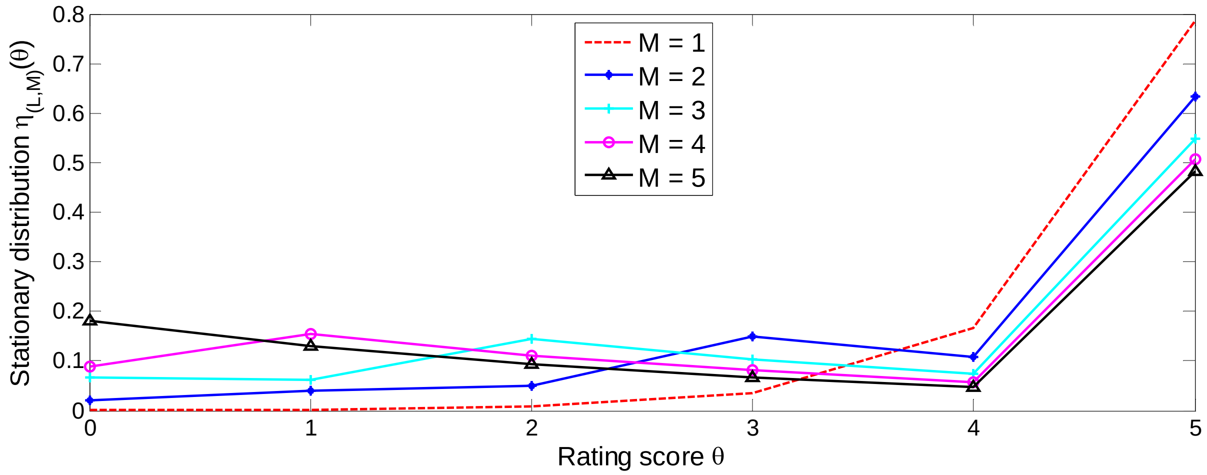
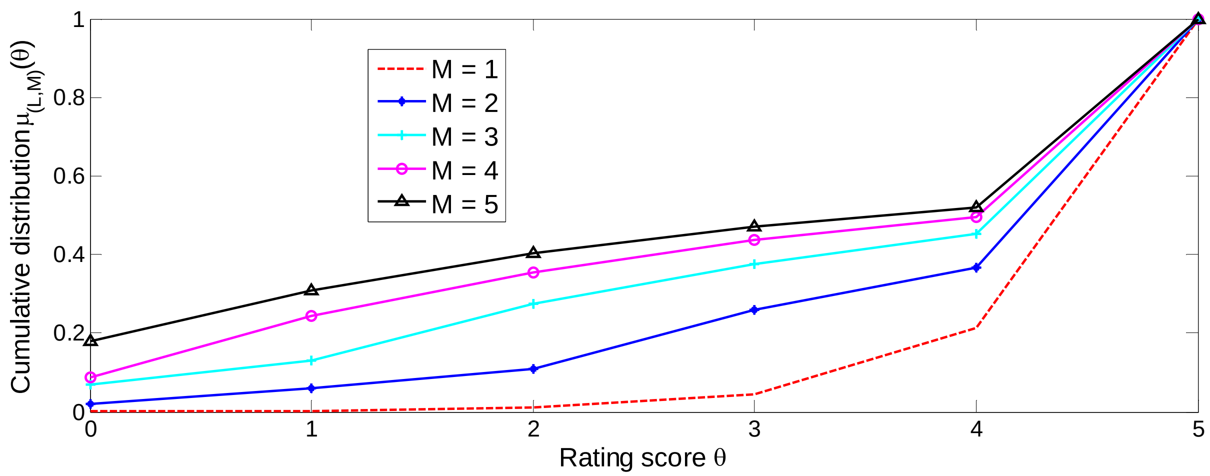


Fig. 4. Optimal social welfare and the optimal recommended strategy of DP_2.



(a)



(b)

Fig. 5. (a) Stationary distribution of rating scores and (b) the cumulative distribution when $L = 5$.

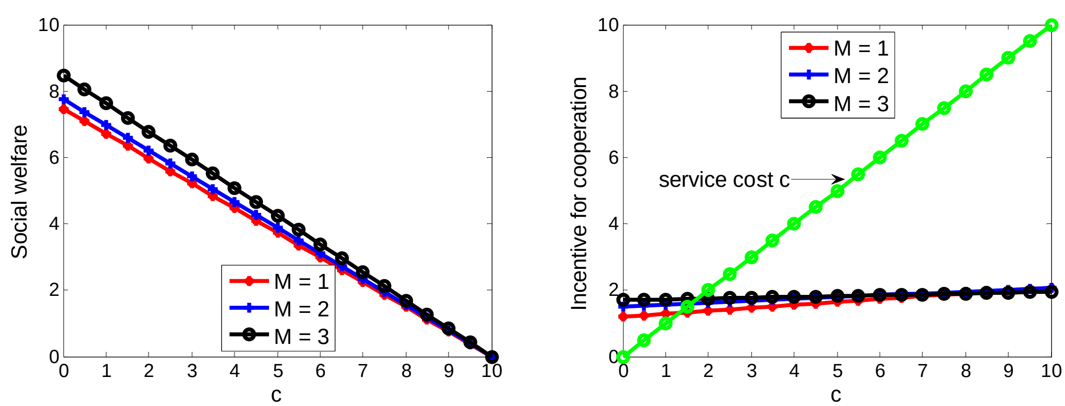


Fig. 6. Social welfare and the incentive for cooperation under recommended strategy σ_L^C when $L = 3$.

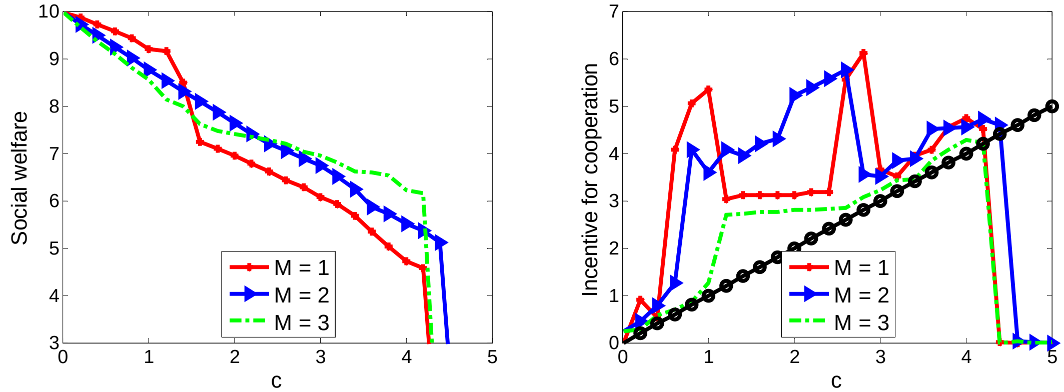


Fig. 7. Social welfare and the incentive for cooperation under the optimal recommended strategy when $L = 3$.

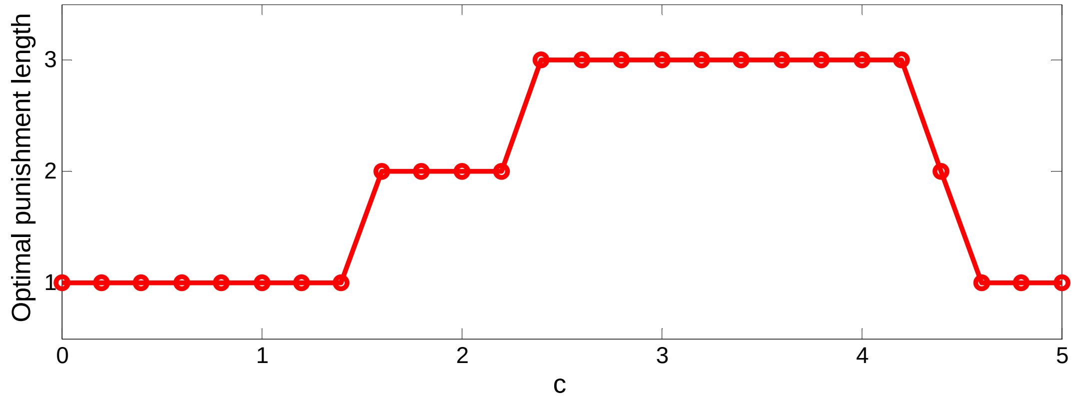


Fig. 8. Optimal punishment length when $L = 3$.

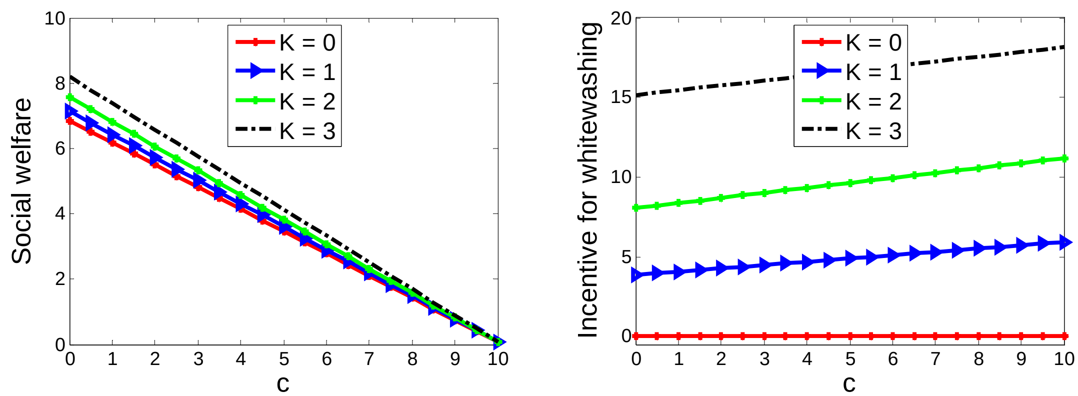


Fig. 9. Social welfare and the incentive for whitewashing under recommended strategy σ_L^C when $L = 3$ and $c_w = 1$.

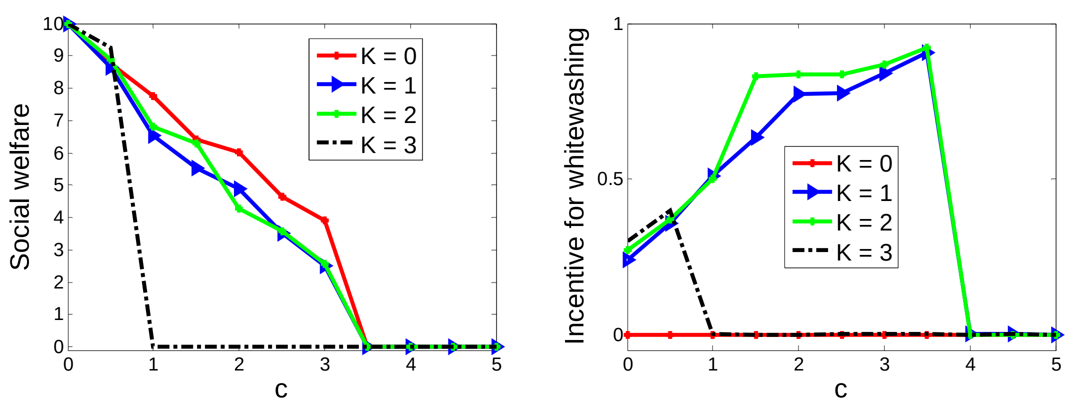


Fig. 10. Social welfare and the incentive for whitewashing under the optimal recommended strategy when $L = 3$ and

$c_w = 1$.

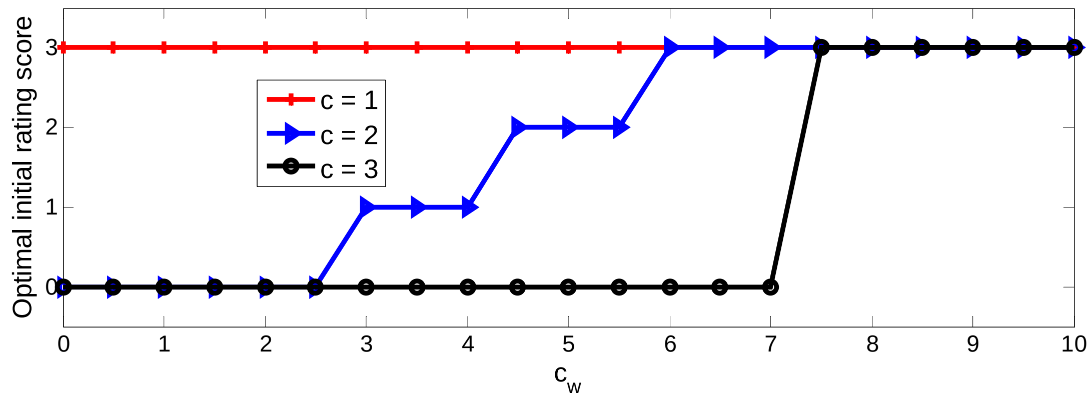
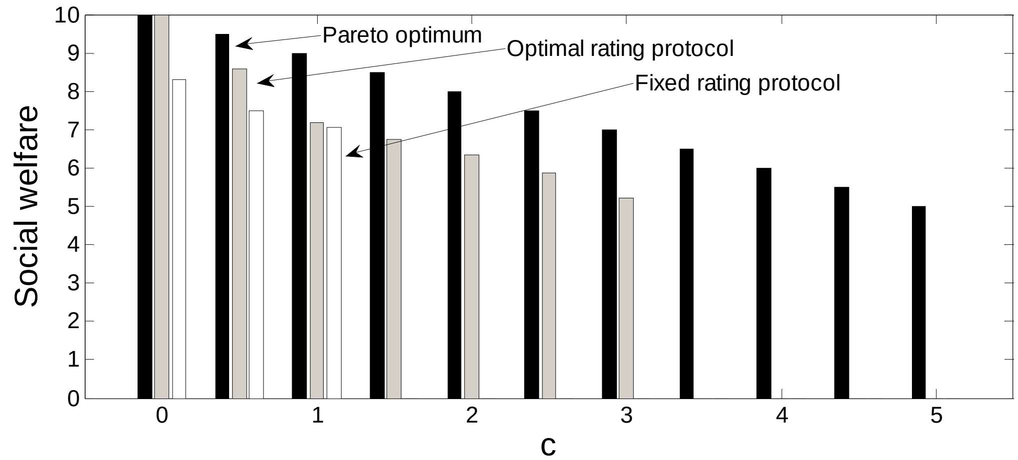
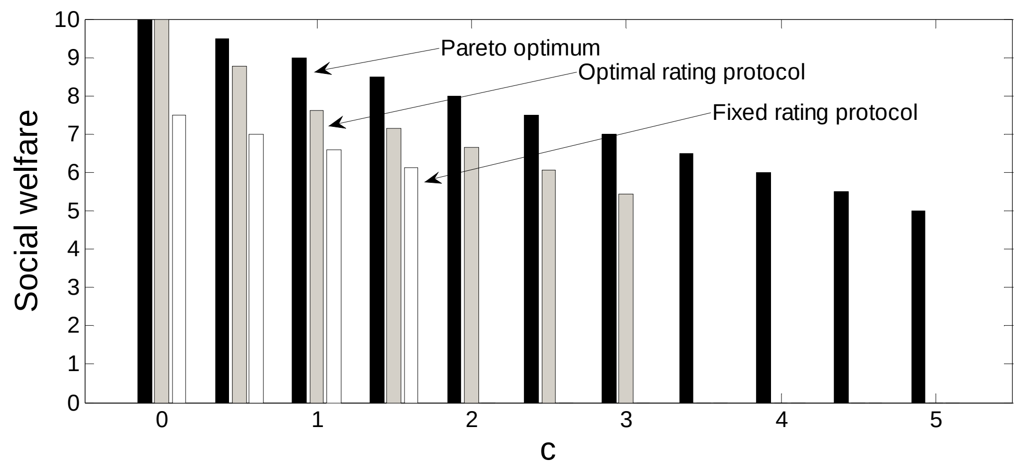


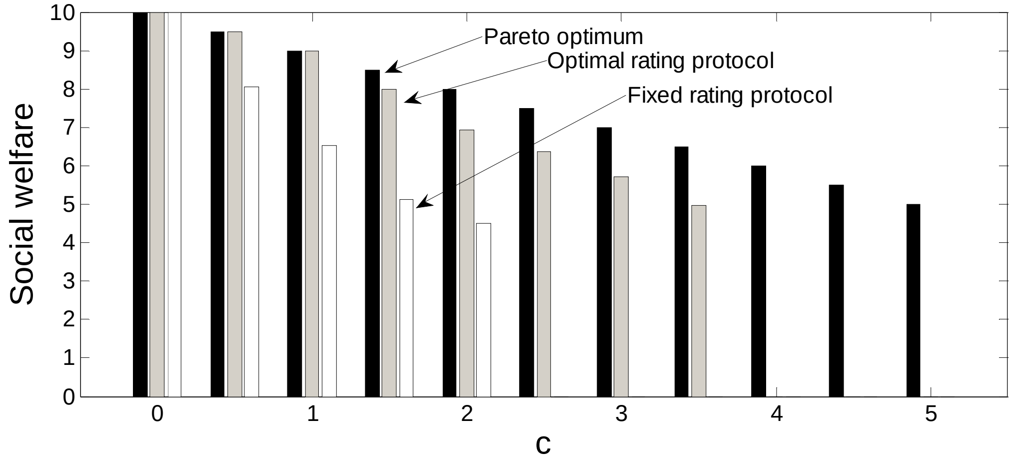
Fig. 11. Optimal initial rating score when $L = 3$.



(a)

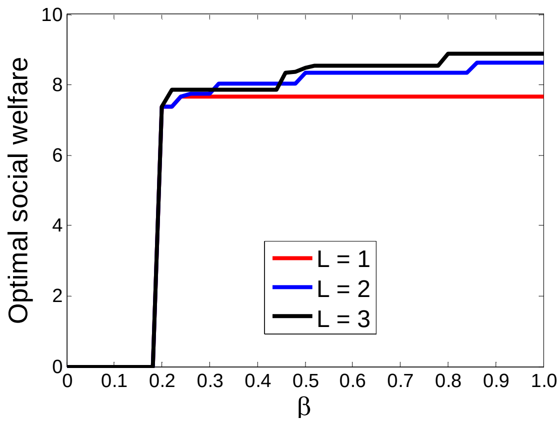


(b)

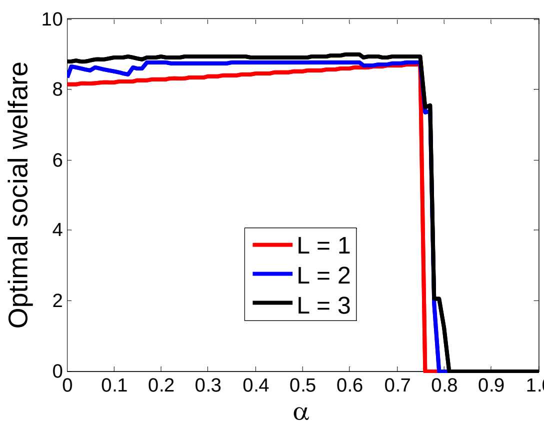


(c)

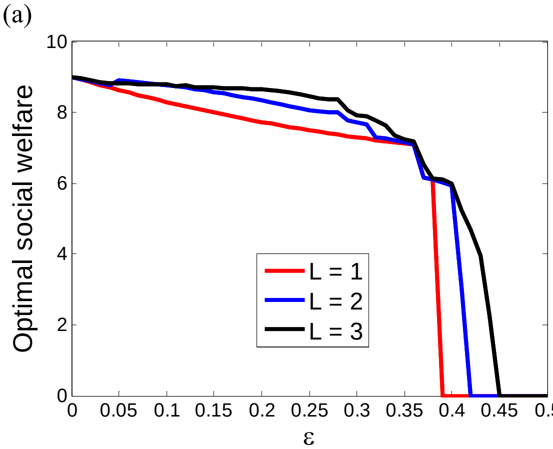
Fig. 12. Performances of the optimal protocol $\left(L, \sigma_L^*\right)$ and the fixed rating protocol $\left(L, \sigma_L^C\right)$ when (a) $L = 1$; (b) $L = 2$; (c) $L = 3$



(a)



(b)



(c)

Fig. 13. Optimal social welfare given L as β, α, and ε vary.

REFERENCES

[1] L. Adamic, J. Zhang, E. Bakshy, M. Ackerman, Knowledge Sharing and Yahoo Answers: Everyone Knows Something, *in 17th Int'l Conf. on World Wide Web (WWW)*, 665 – 674, 2008.

[2] W. Wang, V. Srinivasan and K. Chua, Using Mobile Relays to Prolong the Lifetime of Wireless Sensor Networks, *in 11th Annual Int'l Conf. on Mobile Computing and Networking*, 270 – 283, 2005.

[3] B. Zhao, J. Liu and D. Chiu, A Mathematical Framework for Analyzing Adaptive Incentive Protocols in P2P Networks, *IEEE/ACM Trans. on Networking*, 20(2), 367 – 380, 2012.

[4] K. Ranganathan, M. Pipeanu, A. Sarin and I. Foster, Incentive Mechanisms for Large Collaborative Resource Sharing,

IEEE Int'l Symposium on Cluster Computing and the Grid, 1 – 8, 2004.

[5] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn and S. Moon, I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System, *in 7th ACM SIGCOMM*, 1 – 14, 2007.

[6] A. Blanc, Y. Liu and A. Vahdat, Designing Incentives for Peer-to-Peer Routing, *in IEEE INFOCOM*, 1(1), 374 – 385, 2005.

[7] aN. Miller, P. Resnick, R. Zeckhauser, Eliciting Honest Feedback: The Peer-Prediction Method, *Management Science*, 51(9), 1359 – 1373.

[8] N. Hanaki, A. Peterhansl, P. Dodds, D. Watts, Cooperation in Evolving Social Networks, *Management Science*, 53(7), 1036 – 1050.

[9] B. Awerbuch, B. Patt-Shamir, D. Peleg, M. Tuttle, Collaboration of Untrusting Peers with Changing Interests, *in 5th ACM Conf. on Electronic Commerce*, 112 – 119, 2004.

[10] S. Saroiu, P. K. Gummadi, S. D. Gribble, A Measurement Study of Peer-to-Peer File Sharing Systems, *in Multimedia Computing and Networking*, 156 – 170, 2002.

[11] D. Lopez-Pintado, The Spread of Free-Riding Behavior in a Social Network, *Eastern Economic Journal*, 34, 464 – 479.

[12] D. Bergemann, D. Ozman, Optimal Pricing with Recommender Systems, *in 7th ACM Conf. on Electronic Commerce*, 43 – 51, 2006.

[13] J. K. MacKie-Mason, H. R. Varian, Pricing Congestible Network Resources, *IEEE J. Sel. Areas Commun.*, 13(7), 1141–1149.

[14] V. Vishnumurthy, S. Chandrakumar and E. Gun Sirer, KARMA: A Secure Economic Framework for P2P Resourcing Sharing, *in 3rd Workshop on Economics of P2P Systems*, 2003.

[15] J. Park, M. van der Schaar, Medium Access Control Protocols with Memory, *IEEE/ACM Trans. on Networking*, 18(6), 1921 – 1934.

[16] R. Ma, S. Lee, J. Lui, D. Yau, Incentive and Service Differentiation in P2P Networks: a Game Theoretic Approach, *IEEE/ACM Trans. on Networking*, 14(5), 978 – 991, 2006.

[17] M. Feldman, K. Lai, I. Stoica, J. Chuang, Robust Incentive Techniques for Peer-to-Peer Networks, *in ACM Conf. on Elec. Commerce*, 102 – 111.

[18] A. Habib, J. Chuang, Service Differentiated Peer Selection: an Incentive Mechanism for Peer-to-Peer Media Streaming, *IEEE Trans. on Multimedia*, 8(3), 610 – 621.

[19] P. Resnick, R. Zeckhauser, E. Friedman, K. Kuwabara, Rating System, *Communications of the ACM*, 43(12), 45 – 48.

[20] E. Friedman, P. Resnick, The Social Cost of Cheap Pseudonyms, *Journal of Economics and Management Strategy*, 10(2), 173 – 199.

[21] P. Resnick and R. Zeckhauser, Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System, *Advances in Applied Microeconomics*, 11, 127 – 157.

[22] Y. Zhang and M. van der Schaar, Strategic Learning and Robust Protocol Design for Online Communities with Selfish Users, *UCLA Technical Report*, http://arxiv.org/abs/1108.5514.

[23] C. Dellarocas, Rating Mechanism Design in Online Trading Environments with Pure Moral Hazard, *Information Systems Research*, 16(2), 209 – 230.

[24] C. Dellarocas, How Often Should Rating Mechanisms Update a Trader's Rating Profile, *Information Systems Research*, 17(3), 271 – 285.

[25] R. Jurca, Truthful Reputation Mechanisms for Online Systems, *in 6th Int'l Conf. on Autonomous Agents and Multiagent Systems*, 1(5), 2007.

[26] M. Kandori, Rating protocols and Community Enforcement, *Rev. Economic Studies*, 59(1), 63-80.

[27] M. Okuno-Fujiwara, A. Postlewaite, Rating protocols and random matching games, *Games Econ. Behavior*, 9(1), 79–109.

[28] S. Takahashi, Community Enforcement when Players Observe Partners' Past Play, *Journal of Economic Theory*, 145(1), 42 – 62.

[29] G. Ellison, Cooperation in the Prisoner's Dilemma with Anonymous Random Matching, *The Review of Economic Studies*, 61(3), 567 – 588.

[30] J. Deb, Cooperation and Community Responsibility: A Folk Theorem for Random Matching Games with Names, mimeo.

[31] S. Adlakha, R. Johari, G. Y. Weinstraub, A. Goldsmith, Oblivious Equilibrium for Large-Scale Stochastic Games with Unbounded Costs, *in IEEE Conference on Decision and Control*, 5531 – 5538, 2008.

[32] P. Johnson, D. Levine, W. Pesendorfer, Evolution and Information in a Gift-Giving Game, *J. Econ. Theory*, 100(1), 1 – 21.

[33] D. Kreps, Decision Problems with Expected Utility Criteria, I: Upper and Lower Convergent Utility, *Mathematics of Operations Research*, 2(1), 45 – 53.

[34] D. Kreps, Decision Problems with Expected Utility Criteria, II: Stationary, *Mathematics of Operations Research*, 2(3),

266 – 274.
[35] D. Fudenberg, D. Levin, E. Maskin, “The Folk Theorem with Imperfect Public Information,” *Econometrica*, 62(5), pp. 997 – 1039, 1994.
[36] G. J. Mailath, L. Samuelson, Repeated Games and Reputations: Long-Run Relationships, *Oxford University Press*, 2006.
[37] L. Massoulie, M. Vojnovic, Coupon Replication Systems, *IEEE/ACM Trans. on Networking*, 16(3), 603 – 616, 2005.
[38] P. Whittle, Optimization Over Time, *New York: Wiley*, 1983.