Persuasion for the Long Run

James Best and Daniel Quigley*

February 6, 2020

Abstract

We examine persuasion when the desire for future credibility is the sole source of credibility today. A long-run sender plays a cheap talk game with a sequence of short-run receivers who observe some public record of past feedback about accuracy. We use a geometric approach to provide necessary and sufficient conditions under which long-run incentives can efficiently substitute for ex ante commitment when receivers observe all feedback; these conditions are not met in a large class of settings. We then show that coarse summary statistics of past feedback, like those used online, allow long-run incentives to perfectly substitute for commitment. (JEL C72, C73, D02, D82, D83)

Credible communication often relies on the desire to be believed not only today, but in the future as well. In many settings, like large anonymous markets, these long-run incentives depend in turn on the availability of public records. By publishing feedback about the quality of past communication, records make it possible to punish unsatisfactory advice with future incredulity. Yet such an endogenous source of commitment comes at a cost, as surplus-burning *punishments* must now occur whenever an inaccuracy is published.[1] This creates a trade-off new to persuasion problems: today's

*Best: Carnegie Mellon University Tepper School of Business (e-mail: jabest@andrew.cmu.edu). Quigley: Department of Economics and Exeter College, University of Oxford (e-mail: daniel.quigley@economics.ox.ac.uk). We thank Vince Crawford, Laura Doval, Johannes Hörner, Ian Jewitt, Emir Kamenica, Elliot Lipnowski, Laurent Mathevet, Margaret Meyer, John Hardman Moore, Alessandro Pavan, Doron Ravid, Maryam Saeedi, Ali Shourideh, Steven Tadelis, Ina Taneva, Adrien Vigier, Peyton Young and Bill Zame for their feedback.

[1]Fudenberg and Levine (1994) show a folk theorem does not apply outside long-run relationships.

approach to persuasion must be balanced against its likely effect on future credibility. Clearly, the process for publishing feedback must determine this trade-off. Hence, to analyze persuasion in these settings we must understand the interaction between the individual persuasion problem and the design of the public record system itself.

With the rapid growth of online markets, managing these trade-offs is becoming ever more important. Indeed, platforms such as eBay, Airbnb, and Upwork specifically control participants' public records via coarse ratings. In particular, each promises its vendors a 'badge' if customers' feedback meets the platform's standards. Yet at the same time each hides individual complaints from the customers. Even where individual reviews are available, customers still largely rely on badges.[2] An advantage of such badges is to guide communication: to keep a good record, vendors must meet the platform's standards. But we might worry that such systems weaken long-run incentives by hiding instances of feedback that could be used to punish participants. Perhaps surprisingly, this is not the case: we will see that, by selectively hiding feedback, they allow us to provide incentives with the *reward* of more effective future persuasion instead. Indeed, treating the choice of record as an information design problem, we show simple badge systems can be used to characterize the limits of persuasion when the only source of commitment is repetition and some public record of past advice.

To this end, we develop a model of long-run persuasion where a patient long-run sender ('he') plays a cheap talk game with short-run receivers (each 'she') who arrive sequentially. Each period a payoff-relevant state is independently drawn. The sender observes the state and sends a message to the receiver, who then chooses an action. After the receiver acts, a signal of the state (e.g. feedback) is realized. We say monitoring is perfect when the signal is noiseless, and imperfect otherwise. Repeated play of the game generates a history of messages, actions, and signals. The sender always observes this history. Receivers may observe less: we say the public record is 'complete' if receivers also observe the full history (as is standard in repeated games), and 'incomplete' otherwise. Our focus on short-run receivers ensures the sender can influence actions only via beliefs about the state, allowing for a natural comparison to the commitment benchmark of Kamenica and Gentzkow (2011) (referred to as KG hereafter).

[2] For example, eBay's Top Rated badge requires that sellers avoid too many "*not as described* requests". Importantly, such requests are hidden from public view. Moreover, Nosko and Tadelis (2015) show individual reviews are of limited use and little relied upon; 90% of eBay sellers get $\geqslant 98\%$ good reviews, and $\leqslant 1\%$ of visitors click to see the details behind a rating.

In section 2, we begin with an analysis of complete public records. In this case, we can apply the results of Fudenberg et al. (1990) to demonstrate how long-run incentives fall short even when monitoring is perfect. We show a patient sender can achieve the commitment benchmark if and only if his optimal information structure under commitment is *partitional*, i.e., a deterministic mapping from the state space to the message space. With perfect monitoring, partitional information structures such as full information can be enforced via off-path punishments. However, non-partitional information involves mixing between two different messages in some state. Such mixing requires indifference, which in turn requires (surplus burning) on-path punishments.[3]

We then leverage this observation to characterize the costs of long-run persuasion in terms of the underlying primitives, via a novel geometric approach. To do so we introduce the notion of a *partitional value function*, $\tilde{v}$. Given a prior, $\tilde{v}$ specifies the sender's maximal payoff from any (Bayes-plausible) partitional information structure. In Theorem 1 we establish that long-run incentives are a *perfect* substitute for ex ante commitment in persuasion if and only if $\tilde{v}$ is concave. This analysis yields a result complementary to KG: while they show the benefit from persuasion is given by the difference between the concave envelope of the sender's value function and that function itself, we show the commitment costs of long-run incentives are identified by the difference between the concave envelope of $\tilde{v}$ and $\tilde{v}$ itself.

The geometry of the partitional value function ultimately depends on conflicts of interest over information sharing. For example, there is no such conflict when a committed sender prefers full information irrespective of priors, while conflict is absolute when he never wants to share any information. Yet in both cases $\tilde{v}$ is concave. More generally, Proposition 3 shows $\tilde{v}$ is concave if the state space can be partitioned in such a way that there is no conflict *between* the elements but absolute conflict *within* them. Moreover, it is always optimal (with or without commitment) to communicate over this partition irrespective of priors: the sender says in which element the state lives but nothing more. Hence, even when the analyst (or the sender) have limited information about the receivers' prior beliefs we can still identify the optimal communication strategy — particularly useful in more applied settings where such information is scarce.

Fundamentally, the cost of non-partitional communication results from the receivers'

[3]Mathevet et al. (2019) makes the same initial observation. They then take a different approach from us, studying a reputational-types model in which the sender is committed with some probability.

collective inability to incentivize credible communication with anything other than punishments. From Fudenberg and Levine (1994) we know this problem is only exacerbated when monitoring is imperfect, as even partitional structures will now require on-path punishment — harming sender and receivers alike. This begs a question: is there some form of reward that could be used to provide incentives more efficiently?

To fix ideas, consider a seller advertising products of differing qualities at a fixed price online where customers leave reviews after making a purchase. Suppose that instead of incentivizing via punishments, customers could instead promise to reward the seller by buying products (irrespective of its quality) if the last review was good. Clearly the seller would be better off, but perhaps surprisingly we show this can make the customers better off too: reducing inefficient punishments can more than offset the ex ante cost of providing these (efficient) rewards. Unfortunately, no individual customer would ever knowingly buy a low-quality product. Herein lies the problem.

In section 3 we present our second main insight: by applying information design to the public record, receivers can be persuaded to grant these rewards after all.[4] To do this, we introduce a third-party platform that designs a public record system. Theorem 2 shows how a *simple badge system*, similar to those found online, can eliminate the costs of long-run persuasion. The platform collects feedback from receivers and periodically evaluates it against a set of standards. These standards might, for example, require an online seller to keep the rate of complaints below some threshold. The platform awards the sender with a public badge if and only if these standards are met. While the sender observes all feedback, incoming receivers can *only* observe the badge; even the dates of evaluation are hidden from them.[5]

A desire to keep the badge incentivizes communication likely to meet the (strict) standards. Yet at some private histories the sender has no incentive to be credible. For example, consider the day prior to evaluation: if the feedback has already met the standards at this point then he will certainly lie. The simple badge system solves this problem by pooling these incredible histories with the credible. Indeed, each receiver believes she faces the *average* communication strategy – never knowing when she is

[4]In Sugaya and Wolitzky (2018) hiding history helps long-run firms collude. By contrast, we show how to deal with inefficiencies caused by short-run players: solving the mixing problem of Fudenberg et al. (1990) and the imperfect monitoring problem of Fudenberg and Levine (1994). See section 5.

[5]Similar to Kremer et al. (2014) we model this as hiding information about uncertain arrival time within an evaluation phase. We discuss various foundations for this assumption in section 4.

rewarding the sender at a cost to herself. Under a broad set of conditions, the use of a simple badge system can costlessly substitute for the sender's lack of commitment.

These results demonstrate the scope for review systems to improve the efficiency of persuasion when repetition, along with a public record of accuracy, is the sole source of commitment. This is particularly relevant to the flourishing of online trade.[6] As Tadelis (2016) points out, this success has relied largely on the development of feedback and reputation systems,[7] which have been carefully refined by platforms since their inception and now often take the form of badges. For instance, eBay periodically assesses its sellers and awards a Top Rated badge to those who have attracted disputes on no more than 0.5% of their sales. To avoid disputes, sellers are advised to "describe (each) item accurately". Airbnb provides a Superhost badge that requires accurate communication about availability (among other things). Upwork's badge requires positive feedback 80% of the time, which depends in turn on workers accurately describing their fit for each job.[8] These badges all depend on information hidden from other users. eBay hides the rate of disputes and Airbnb hides the ratings given by individuals (revealing only an aggregate). Finally, both eBay and Upwork provide vendors with portals to *privately* track their feedback and to see if they are in danger of losing their badge.

One might worry that platforms may face issues committing to a simple badge system. In section 4 we discuss this issue in light of an alternative (sender-blind) review system. In doing so, we also formally identify the theoretical and applied differences between simple badge systems and systems that 'reuse punishments', as in Abreu et al. (1991) and Fuchs (2007). We also extend our analysis to evaluate review systems that cannot hide evaluation dates or acquire feedback, and those that face other forms of moral hazard. Finally, we discuss related literature in section 5 and conclude in section 6. Proofs are contained in the Appendix.

[6] Having grown from 0.5% to 10% of retail sales in the past two decades (US Bureau of the Census).

[7] Hui et al. (2016) find that eBay's Top Rated badge both confers significant reputational advantages to sellers and motivates those in danger of losing the badge to improve their behavior.

[8] Of course, other forms of moral hazard also affect online trade. We discuss how badges can efficiently tackle these problems too (see section 4.4).

1 Model

We consider a general repeated cheap talk game between a long-lived sender with discount rate δ (he) and a sequence of short-lived receivers (each she). At each time $t = 1, 2, \ldots$, the current state of the world θ_t is drawn randomly from a common prior μ_0 over a finite set Θ; the probability of θ is $\mu_0(\theta)$. The sender privately observes θ_t and sends a message $m_t \in M$ to the receiver, who chooses an action a_t from a compact set A. After she takes her action a noisy signal $\omega_t \in \Omega$ is drawn from the conditional distribution $p(\omega_t \mid \theta_t, a_t)$, where (θ_t, ω_t) is history-independent for given a_t. Hence, one can interpret $(m_\tau, a_\tau, \omega_\tau)$ as feedback about the accuracy of communication. We say monitoring is **perfect** if ω_t perfectly identifies θ_t for any $a_t \in A$ and **imperfect** otherwise. The sender and receiver have respective stage payoffs $v(a)$ and $u(a, \theta, \omega)$ which are continuous in a. We assume the sender cares only about a so his short-term incentive is to say whatever induces his favored action. This is for ease of exposition: all results (with the exception of Corollary 3) hold for state-dependent preferences.

The sender observes past $\theta^t = (\theta_1, \ldots, \theta_{t-1})$ and the **history of feedback**, $\underline{h}_t = (m_\tau, a_\tau, \omega_\tau)_{\tau=1}^{t-1}$, with $\underline{h}_1 = \varnothing$. Denote the sender's (private) history by $h_t = (\theta^t, \underline{h}_t)$, and let the corresponding sets of histories at t be $\underline{\mathcal{H}}_t$ and $\mathcal{H}_t$, respectively. The receiver at t only observes a public history, r_t. We allow the public history to be a potentially coarse measure of past feedback. It is determined by a sequence of functions, $r_t : \mathcal{H}_t \to \mathcal{R}_t, t = 1, 2, \ldots$, where each $\mathcal{R}_t$ contains the (finite) set of possible public histories at t. The standard case of $r_t = \underline{h}_t$ is called a **complete public record**; otherwise, it is **incomplete**. We also allow players access to a public randomization device. Strategies and belief systems have the standard definitions.

Throughout the paper, we illustrate our main results with the following example:

Example 1 *A long-lived seller serves a sequence of anonymous customers. Each period, the seller has a product of quality $\theta \in \{l < 0, h = 1\}$, where $\Pr[\theta = h] = \mu_0$. He makes a claim $m \in \{$'like new', 'used'$\}$. If a customer chooses to* 'Refuse' *the product, she and the seller get* 0*. If she chooses* 'Buy' *she gets a payoff θ and the seller gets* 1*. The conditional distribution of the signal $\omega \in \{b, g, \varnothing\}$ is $p(g \mid h, Buy) = \Pr(b \mid l, Buy) = p > \frac{1}{2}$ and $p(\varnothing \mid h, Refuse) = 1$. Hence, g, b and $\varnothing$ can be interpreted as good, bad, and no review respectively.*

Benchmarks for Long-Run Persuasion

In any stage game, the sender's strategy induces an **information structure**: a distribution $\lambda \in \Delta\Delta\Theta$ over the receiver's posterior beliefs, where $\lambda(\mu)$ and $\lambda(\mu \mid \theta)$ are respectively the unconditional and conditional probabilities that the receiver's induced posterior is μ.[9] When the sender can commit to his communication strategy, KG shows it is without loss of generality to consider his strategy as a choice of information structure λ from a *Bayes-plausible* set, $\Lambda(\mu_0) = \{\lambda : \mathbb{E}_\lambda[\mu] = \mu_0\}$. Where obvious, we drop the dependence of Λ on μ_0.

Similarly, stage payoffs can be written in terms of receiver beliefs. With a slight abuse, let the sender's payoff given posterior μ be $v(\mu)$, and his ex ante payoff from λ be $V(\lambda) := \mathbb{E}_\lambda[v(\mu)]$. Let $u(\mu)$, $U(\lambda)$ be the analogous receiver payoffs. The persuasion literature focuses on the sender's maximal commitment payoff, or **Stackelberg payoff**:

$$\hat{v}(\mu_0) := \max_{\lambda \in \Lambda} V(\lambda). \tag{KG}$$

As KG shows, this payoff is given by the concave envelope of v: $\hat{v} = \text{cav } v$.[10] Let **Stackelberg persuasion**, $\hat{\lambda} \in \hat{\Lambda}(\mu_0)$, be any λ solving (KG).[11]

As receiver payoffs matter too, we also consider the commitment payoff set $\mathcal{CS}$:

$$\mathcal{CS} := \left\{ (u,v) : u = U(\lambda), v = V(\lambda) \text{ , for some } \lambda \in \Lambda(\mu_0) \right\}.$$

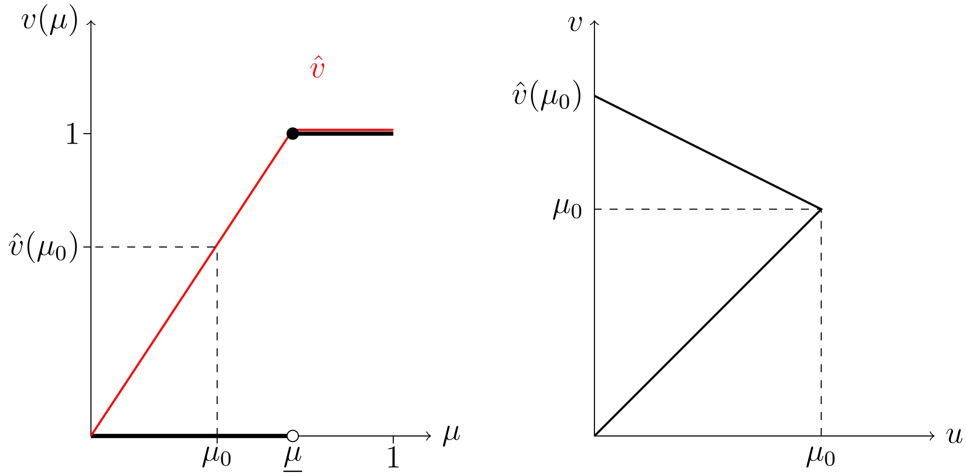
Figure 1 illustrates these concepts for Example 1. The customer buys if and only if her posterior is $\mu \geqslant \underline{\mu} := \frac{|l|}{1+|l|}$. This gives v as the thick black line in Fig. 1a, $\hat{v}$ as the red line, and $\hat{v}(\mu_0) = \frac{\mu_0}{\underline{\mu}}$. Notice that if $\mu_0 \geqslant \underline{\mu}$ then $\hat{\lambda}$ will be uninformative. Otherwise, $\hat{\lambda}$ splits posteriors between 0 and $\underline{\mu}$. This is equivalent to the seller reporting "like new" with conditional probabilities $\lambda(\underline{\mu}|\theta = h) = 1$ and $\lambda(\underline{\mu}|\theta = l) = \frac{\mu_0}{(1-\mu_0)|l|}$. Of course, the customer optimum is full information, i.e. truth telling, which implements (μ_0, μ_0) in 1b. The line connecting it to $(0, \hat{v}(\mu_0))$ gives the Pareto frontier of the commitment set.

[9] We consider only those λ with finite support (denoted supp λ). This is w.l.o.g.: see Lemma 1.

[10] We adopt the standard notation that cav f denotes the concavification of a function f.

[11] Strictly, $V(\lambda)$ ought to be an equilibrium correspondence. Since we care about feasible equilibrium payoffs, we adopt sender-preferred equilibrium selection when discussing his optimal payoff.

Figure 1: Stackelberg Persuasion in Example 1



(a) Stackelberg Persuasion (b) Commitment Payoff Set, $\mathcal{CS}$

Finally, we assume throughout that the sender's optimal payoffs cannot be achieved with informative cheap talk in the one-shot game. Lipnowski and Ravid (2019) show the sender's maximal equilibrium payoff in the one-shot game is characterized by the quasi-concave envelope of v, denoted $\bar{v}(\mu)$.[12] In their notation, our assumption is:[13]

Assumption 1 *If $v(\mu) < \hat{v}(\mu)$, then $\bar{v}(\mu) < \hat{v}(\mu)$.*

2 Persuasion with Complete Public Records

We begin by investigating how well repetition substitutes for ex ante commitment in persuasion in the standard public monitoring setting where public records are complete.

2.1 Preliminaries

While our setting features repeated play and no exogenous commitment, we show an analogue of KG's belief-based approach still applies. Such an approach is convenient:

[12] The quasi-concave envelope is the smallest quasi-concave function everywhere greater than $v(\mu)$.

[13] An earlier version of this paper (Best and Quigley, 2017) shows that functions violating Assumption 1 are non-generic. Similarly, Corollary 3 in Lipnowski and Ravid (2019) shows that cheap talk cannot substitute for informative Stackelberg persuasion in the finite action case.

it facilitates comparison with Stackelberg persuasion and will allow us to geometrically characterize the costs of long-run persuasion.

Consider restricting the sender's message space to the set of recommended posterior beliefs: $m_t \in \Delta\Theta$ for all t. Then, the sender's (mixed) strategy is just a sequence of functions $\sigma_t : \underline{\mathcal{H}}_t \to \Delta\Delta\Theta$. Similarly, receivers' posteriors can be described by a sequence of functions $\mu_t : \Delta\Theta \times \underline{\mathcal{H}}_t \to \Delta\Theta$. While in general there is no reason for receivers' beliefs to coincide with m_t or for the sender's chosen λ_t to even be Bayes-plausible, we can nonetheless define a class of equilibria in which these additional restrictions hold:

Definition 1 *A **direct equilibrium** (σ, μ) is a Perfect Public Equilibrium (PPE) in which (i) $\sigma_t = \lambda_t \in \Lambda(\mu_0)$, and (ii) $\mu_t = m_t \in \Delta\Theta$ for all on-path $(m_t, \underline{h}_t)$.*

As in KG, on the path of a direct equilibrium λ_t is a choice of Bayes-plausible information structure. Unlike the standard environment however, our sender's language is not as rich as his private information: he knows the entire sequence $\theta_1, ..., \theta_t$ but can communicate only about θ_t. Moreover his strategy only depends on the public record. Nonetheless, we show a belief-based approach is still without loss:[14]

Lemma 1 *For any Perfect Bayesian Equilibrium of the game, there is a direct equilibrium that induces the same on-path distribution over beliefs, actions and payoffs, $\forall t$. Moreover, $|\text{supp } \lambda_t| \leqslant \min\{|\Theta|, |A|\}$ at each $\underline{h}_t$ without loss.*

Lemma 1 allows us to directly compare the sender's payoffs in long-run and Stackelberg persuasion. In Stackelberg persuasion the sender can commit, so the only constraint in (KG) is Bayes plausibility. In long-run persuasion a sender-optimal equilibrium solves

$$\max_{(\lambda_t)_{t=1}^{\infty}} (1-\delta) V(\lambda_1) + \delta \mathbb{E}_{\lambda_1}\left[V_2^{LP}(h_1)\right] \qquad \text{subject to}$$

$$V_t^{LP}(h_t) = (1-\delta)V(\lambda_t(h_t)) + \delta \mathbb{E}_{\lambda_t}\left[V_{t+1}^{LP}(h_{t+1})\right], \qquad \lambda_t \in \Lambda(\mu_0), \forall t, h_t \in \mathcal{H}_t, \tag{1}$$

$$V_t^{LP}(h_t) \geqslant (1-\delta)V(\lambda_t') + \delta \mathbb{E}_{\lambda_t'}\left[V_{t+1}^{LP}(h_{t+1})\right], \qquad \forall t, h_t \in \mathcal{H}_t, \lambda_t' \in \Delta\Delta\Theta,$$

where V_t^{LP} represents discounted average continuation payoffs. The first set of constraints requires Bayes plausibility for all on-path histories (see KG). The second set

[14] Lipnowski and Ravid (2019) have a similar result in the context of one-shot cheap talk.

are the incentive constraints resulting from the lack of ex ante commitment: the sender must not wish to deviate from his equilibrium information provision. Trivially, (1) is bounded above by $\hat{v}(\mu_0)$. A more meaningful bound can be established with the following concept. Given an information structure λ, define

$$\underline{\mu}_\theta(\lambda) := \arg\min_{\mu \in \operatorname{supp} \lambda|\theta} v(\mu)$$

as the sender's *worst message in state* θ, with a corresponding *worst stage payoff* $\underline{v}_\theta(\lambda)$. Applying a result of Fudenberg et al. (1990) bounds the sender's payoffs as follows:[15]

Proposition 1 ***(Fudenberg et al., 1990)*** *The sender's optimal continuation payoff in any long-run persuasion game, V_t^{LP}, is bounded above by*

$$\max_{\lambda \in \Lambda(\mu_0)} \sum \mu_0(\theta)\, \underline{v}_\theta(\lambda). \tag{2}$$

There is $\underline{\delta} < 1$ s.t. (2) is attainable in equilibrium if $\delta \in [\underline{\delta}, 1)$ and monitoring is perfect.

The sender's equilibrium payoffs never exceed those he would get if he could commit to an information structure but was only able to receive the *worst* payoff in each state. In order to persuade any receiver, the sender's equilibrium information structure must be credible: he must prefer it to any deviation. Hence if his optimal information structure involves mixing between messages in some state θ, he must be indifferent between them in that state. In the sender's best equilibrium, this indifference requires that each message m induces an on-path punishment $v(m) - \underline{v}_\theta(\lambda)$. This exactly offsets the stage gains of that message relative to the worst — pinning down the upper-bound.

Equation (2) can therefore be used to re-frame the dynamic problem (1) as a static costly persuasion problem, whose solution is given by the following function of μ_0:

$$V^{LP}(\mu_0) = \max_{\lambda \in \Lambda(\mu_0)} V(\lambda) - C(\lambda), \tag{LP}$$

where $C(\lambda) := \mathbb{E}\left[v(m) - \underline{v}_\theta(\lambda)\right] \geqslant 0$. $C(\lambda)$ can be interpreted as the expected punishment cost required to support λ at the optimum. Figure 2a illustrates this for Example 1 with $\mu_0 < \underline{\mu}$: the equilibrium payoff set is a strict subset of $\mathcal{CS}$. Even under perfect

[15] Since it is a direct application, we omit the proof. Mathevet et al. (2019) also identifies this bound.

monitoring, the sender can do no better than an equilibrium in which he provides full information on-path (supported by the off-path threat of a babbling continuation).

2.2 Costless Long-Run Persuasion

Equation (LP) gives an upper bound on payoffs from long-run persuasion. As this bound holds regardless of the monitoring structure, it provides a best case for repetition as a substitute for commitment. We first analyze this best case, assuming in this and the next section that monitoring is perfect and $\delta \geqslant \underline{\delta}$. In this setting, we say that **the cost of long-run persuasion** at μ_0 is $\hat{v}(\mu_0) - V^{LP}(\mu_0)$.

We show in Proposition 2 that (LP) is equal to (KG) if and only if Stackelberg persuasion is **partitional**. Let $\mathcal{P} = \{P_1, ..., P_k\}$ denote a partition of Θ and B_Θ the set of all such partitions. An information structure is said to be partitional if it is generated by a pure communication strategy, $\sigma^{\mathcal{P}} : \Theta \to \Delta\Theta$. Each such strategy represents a simple form of communication: messages can be interpreted as claims "θ is in P_i" for some associated partition $\mathcal{P}$. Of course, given any prior, communication over a partition $\mathcal{P}$ induces an information structure. We denote this by $\lambda^{\mathcal{P}}(\mu_0)$ and its associated payoff by $v^{\mathcal{P}}(\mu_0)$. Importantly, such structures require no on-path punishment.

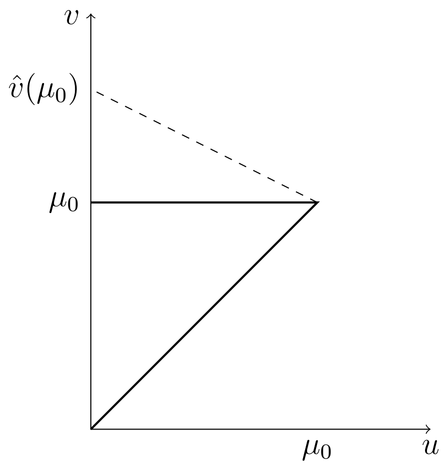
Proposition 2 *$V^{LP}(\mu_0) = \hat{v}(\mu_0)$ if and only if $\exists \mathcal{P} \in B_\Theta$ such that $\lambda^{\mathcal{P}}(\mu_0) = \hat{\lambda}(\mu_0)$.*

The proof is simple: the 'if' direction follows immediately from Proposition 1; the 'only if' direction from Assumption 1 (which ensures messages in supp $\hat{\lambda}$ have different stage payoffs). Ultimately, Proposition 2 only relates the costs of long run persuasion to the endogenous object $\hat{\lambda}$. Yet this intermediate result helps us characterize long-run persuasion in terms of the primitives of the persuasion problem directly.

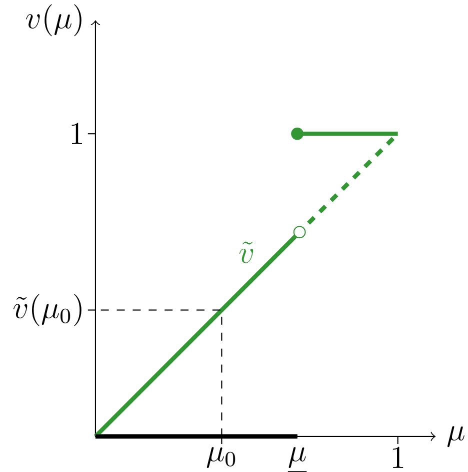
2.3 The Geometry of Long-Run Persuasion

We now establish an equivalence between the performance of repetition as a commitment substitute and the geometry of two functions. This allows us to provide clean necessary and sufficient conditions on preferences under which long-run persuasion can be said to *perfectly substitute* for Stackelberg persuasion. We build on these results to a) bound the

Figure 2: Long-Run Persuasion in Example 1



(a) LRP Payoff Set $p = 1$



(b) Partitional Value Function

costs of long-run persuasion, b) establish when concavification methods can determine the nature of long-run communication, and c) establish when we can restrict attention to a single simple form of communication. First we define a simple auxiliary function:

Definition 2 *The* ***partitional value function*** *is* $\tilde{v} : \Delta\Theta \to \mathbb{R}$*, where*

$$\tilde{v}(\mu) = \max_{\mathcal{P} \in B_\Theta} v^{\mathcal{P}}(\mu). \tag{3}$$

The partitional value function describes the highest payoff a committed sender can earn from partitional information for each prior. Figure 2b illustrates $\tilde{v}$ in the context of Example 1. Notice that $\tilde{v}$ is easily calculated: it is the maximum of finitely many functions, which are each just the expectation of $v(\mu)$ across subsets of Θ. We now show how the geometry of $\tilde{v}$ characterizes the costs of long-run persuasion.

Theorem 1 *For any long-run persuasion game, these statements are equivalent:*

1. $V^{LP} = \hat{v}$;

2. $\tilde{v}$ *is concave;*

3. V^{LP} *is concave.*

Theorem 1 tells us long-run persuasion is costless for any common prior if and only if $\tilde{v}$ is concave. The equivalence follows from two observations: first, since silence is itself trivially partitional and since (3) is a constrained version of (KG), $\tilde{v}$ must always lie in $[v, \hat{v}]$; second, $\hat{v}$ is by definition the smallest concave function that majorizes v. Then, if $\tilde{v}$ is concave it must everywhere equal $\hat{v}$. Hence, long-run persuasion can always attain the Stackelberg payoffs. However, if $\tilde{v}$ is not concave there must be some μ_0 for which $\tilde{v} < \hat{v}$. The consequences for V^{LP} then follow swiftly from Proposition 2.

When $\tilde{v}$ is not concave long-run persuasion is an imperfect substitute for commitment: $V^{LP} \not\equiv \hat{v}$. The geometry of $\tilde{v}$ also allows us to identify at exactly which beliefs these costs will be borne. Moreover, it provides upper bounds on these costs.

Corollary 1 *Given a prior μ_0, $V^{LP}(\mu_0) < \hat{v}(\mu_0)$ if and only if*

$$\tilde{v}(\mu_0) < \mathit{cav}\ \tilde{v}(\mu_0). \tag{4}$$

Moreover, the cost of long-run persuasion is bounded by $\mathit{cav}\ \tilde{v} - \tilde{v} \geqslant \hat{v} - V^{LP} \geqslant 0$.

Hence, the costs of long-run persuasion are identified by comparison of the partitional value function with its own concavification. Here the long-run sender bears costs if and only if cav $\tilde{v} > \tilde{v}$. This offers a complement to KG, who show the sender benefits from persuasion if and only if cav $v > v$. Notice that the same geometry also bounds these costs — providing an immediate worst-case analysis when it would otherwise be difficult to solve for V^{LP}.

While concavification methods therefore bound the costs of long-run persuasion, Theorem 1 cautions that solving for V^{LP} is hard. As Gentzkow and Kamenica (2014) show, persuasion with a posterior-separable cost function is still amenable to concavification methods. Yet in problem (LP), costs are not separable in posteriors: $v(m) - \underline{v}_\theta(\lambda)$ depends both on m and $\underline{\mu}_\theta$. Indeed, Theorem 1 tells us that concavification methods can hope to solve (LP) only when long-run persuasion is *always* costless:

Corollary 2 *V^{LP} can be identified by concavification methods if and only if $\tilde{v} = \mathit{cav}\ \tilde{v}$.*

The geometric approach of Theorem 1 is particularly useful when applied to specific persuasion games. For instance, Corollary 3 below shows that long-run persuasion is bound to be costly in any problem where the receiver chooses from a finite set of actions:

Corollary 3 *Suppose A is finite, the sender's payoffs depend only on a, and the sender benefits from Stackelberg persuasion for some μ_0. Generically, $V^{LP} \neq \hat{v}$.*

The argument generalizes an observation in Figure 2b. As we move prior beliefs outside the region where an uninformed receiver would choose the sender's preferred action, all partitional information structures must induce a less favorable action with a probability bounded away from 0.[16] As a result, $\tilde{v}$ fails to be continuous — and therefore cannot be concave. From Theorem 1, we must have $V^{LP} \neq \tilde{v}$.

Despite this negative result, there are many settings where long-run persuasion is costless. Indeed, take the following trivial cases where $\hat{v}$ is supported by truth or silence.

Corollary 4 *$\tilde{v} = cav\ \tilde{v}$ if either of the following holds for a sender with commitment:*

1. *There is no information the sender would share at any prior.*

2. *The sender (weakly) prefers to share all information at all priors.*

The first claim is obvious: the sender can never benefit from persuasion when v is concave. While the second claim might appear more surprising, one need only consider that the expected value of using the full-information partition is linear.[17] These examples feature very clear extremes in the degree of alignment between sender and receiver over information sharing. At the first extreme, there is an absolute conflict of interest: after all, the sender would never share information. At the other, there is no such conflict. Yet in both cases $\tilde{v}$ is concave. This connection is not accidental: more generally, $\tilde{v}$ continues to be concave whenever the state space can be partitioned so that there is no conflict between the elements but absolute conflict within them.

Proposition 3 *There exists a partition $\mathcal{P}$ such that $v^{\mathcal{P}} = \tilde{v} = \hat{v}$ iff:*

[16]We rule out non-generic cases where the receiver only takes actions on a knife-edge set of beliefs. Moreover, Assumption 1 ensures the region on which the sender secures his best payoff is convex.

[17]Note, if v is convex then $\tilde{v}$ is convex. In turn, this implies $\tilde{v} = \text{cav}\ \tilde{v}$. The converses do not hold.

(i) The sender always prefers $\lambda^{\mathcal{P}}$ over silence: $v^{\mathcal{P}} \geqslant v$.

(ii) v is concave on ΔP_i, $i = 1, \ldots, k$.

Proposition 3 gives necessary and sufficient conditions under which communication over a single partition always supports Stackelberg payoffs. These conditions hold whenever the conflict over information sharing can be described in terms of a dichotomy. Interests are aligned over sharing *which element* of a partition arises (from *(i)*, both always prefer sharing $P_i \in \mathcal{P}$ over no communication at all). However, as implied by the concavity in *(ii)*, there is absolute conflict over sharing information *within each element*. By contrast, when $\tilde{v}$ is not concave no such dichotomy exists and long-run incentives are an imperfect substitute for ex ante commitment. For instance, in the setting of Corollary 3 any partition must fail at least one of the conditions of Proposition 3 for some priors: any informative partition fails condition *(i)* at priors inducing the sender's preferred action, while no-information fails (ii) at priors that do not.

Proposition 3 also allows us to identify the optimal communication strategy even when we (or even the sender) have limited information about receivers' beliefs. This may be particularly useful in applied settings where such information is sparse. Moreover, it tells us when the support of equilibrium messages (the equilibrium vocabulary) is stable even when priors are not. Finally, this result also shows when we can restrict attention to communication over a single partition in Stackelberg persuasion too

2.4 Imperfect Monitoring

Under imperfect monitoring one could characterize payoffs by applying the results of Fudenberg and Levine (1994). Rather than undertake this exercise, we will merely note how imperfect monitoring exacerbates the problems identified above. In particular, the set of equilibrium payoffs for all players (weakly) contracts. For instance, when monitoring is imperfect costly on-path punishment can be required to support partitional information structures. This is bad news for all players. Unlike before, now even the receiver-preferred equilibrium (full information) involves inefficient on-path punishment.

Example 1 highlights the intuition. Consider an equilibrium in which the sender provides full information whenever possible (i.e., when not being punished). Now punishments can depend only on m and ω. To prevent him from mis-selling (i.e., announcing "like new" when $\theta = l$), the seller must be punished for a bad review ($\omega = b$). If he did mis-sell, such a review would occur with probability p. Hence the minimal punishment cost, say c, needed to support full information satisfies the indifference condition:

$$1 = \delta pc. \tag{5}$$

Unlike perfect monitoring the seller now also suffers a bad review when $\theta = h$ and thus pays c with probability $1 - p$. Since this is true at *every* history where his advice is credible, his payoff is bounded above by $\mu_0 - \mu_0 \frac{1-p}{p} = \mu_0 \frac{2p-1}{p}$. Notice customers' preferred equilibrium now gives them average payoffs less than μ_0: *both sides* of the market bear the costs of moving to inefficient babbling phases after bad reviews.

3 Persuasion with Incomplete Public Records

It is natural to conjecture that by providing receivers less information on which to base their punishments, incomplete public records could only exacerbate the inefficiencies identified in section 2. After all, we have just seen how imperfect monitoring contracts the set of equilibrium payoffs. Indeed, in the limiting case of the least informative record only stage Nash equilibria can be sustained. And yet in many markets coarse summaries such as star ratings, badges, or certificates, are adopted over more informative alternatives. In this section, we demonstrate the scope for such coarse records to increase the equilibrium payoff set via an information design approach. We introduce a third-party designer (the "platform") who can commit to rules that determine what incoming receivers observe about $\underline{h}_t$. But first, we provide an example that illustrates a simple insight: coarse records can improve payoffs for sender and receivers alike.

3.1 Pooling Histories to Persuade

Consider again Example 1 with imperfect monitoring. We saw how inefficiencies arose because (surplus-burning) punishments were required to support even full information. But what if customers could collectively promise rewards — such as buying even 'used' products in the future — for earning good reviews today? These rewards would act as transfers from future customers to the seller. Hence they might offer scope to improve payoffs. With a complete public record, such promises are not an equilibrium: future customers would always renege on the promise. Our insight is that when records are coarse, customers may instead be 'persuaded' to provide these rewards.

Consider a simple public record: at $t \in \{1, 2\}$, the record remains completely empty. From $t \geqslant 3$ onward, the record consists of a rating $r \in \{A, B, C\}$ for the first two periods along with detailed feedback $(m_\tau, a_\tau, \omega_\tau)$ on all subsequent interactions $\tau = 3, \ldots, t-1$. The seller earns an "A" so long as feedback was good or neutral at $t = 1$, a "B" if feedback was bad at $t = 1$ but good/neutral at $t = 2$, and a "C" if both periods went badly. With this simple change to the game, the equilibrium payoff set strictly expands:

Proposition 4 *If receivers observe an incomplete public history, average equilibrium payoffs of sender and receiver can exceed the bounds of Fudenberg and Levine (1994).*

To prove this we construct a simple equilibrium in which the second customer plays the role of reward giver. The equilibrium has the following features: at $t = 3$, if the seller has earned an A, play immediately continues on to the (uniquely efficient) full information equilibrium from section 2.4. Otherwise, there is a costly phase of babbling whose length decreases in the grade. After any punishments have been served, play again reverts to the full information equilibrium.

The seller has an incentive to be truthful at $t = 1$ (as he pursues an A grade) and also at $t = 2$, *if* his first interaction went badly (to avoid a C). However, if he avoided bad feedback at $t = 1$ he strictly prefers to lie to the second customer (he will earn an A regardless). This "free lie" is the reward associated with earning an A. Coarse records are crucial here: if the second customer could observe the period 1 feedback, she would be able to infer when she was being lied to. But with coarse feedback, she can be persuaded the sender is credible enough for a range of parameters.[18]

Relative to full information, the new equilibrium directly rewards the seller for avoiding a bad review in period 1. Hence, with probability $1 - \mu_0(1 - p)$, he earns an additional revenue of $1 - \mu_0$ from the second customer and escapes any punishment for a bad review at $t = 2$. As we have seen in section 2.4, the latter is worth $\mu_0 \frac{1-p}{p}$ to him. In fact, the benefits are even larger: we show that his equilibrium payoff increases by the

[18]Notice, the sender conditions his strategy on his private history so this is not a Perfect Public Equilibrium — in any such equilibrium the Fudenberg and Levine (1994) bound would apply.

(discounted) sum of these two terms, whether his first review is good, neutral or bad! Since the reward is its own incentive for honesty at $t = 1$, the punishment associated with a B grade no longer needs to be so harsh. Indeed, the efficient reward replaces the inefficient punishment one-for-one. In fact, customers can be made better off too. While the second customer is effectively giving away a transfer, all customers benefit from a reduction in costly on-path punishments. The proof shows the latter effect can dominate the former.

Before turning to our main result, we make a brief observation. Suppose we had instead given a reward in period 3 if the seller avoided bad reviews on both days 1 and 2. For simplicity, suppose only the second review is hidden from customer 3. In this case, similar arguments imply the seller only expects to actually earn the reward with probability $(1-\mu_0(1-p))^2$. As before, his continuation payoff at $t = 2$ can be improved by $1-\mu_0+\mu_0\frac{1-p}{p}$, discounted once. By backward induction, the full gain carries forward to the first period, discounted once more. This suggests that the scope for rewards to improve payoffs could be large when the sender is patient. We turn to this next.

3.2 The Commitment Set by Design

Many online platforms, including eBay, Airbnb and Upwork, provide coarse records of market participants' past outcomes, awarding badges to those they deem to be sufficiently well-behaved. These badge systems share several key features. They monitor the participant during *evaluation phases*, collecting reviews and complaints about misleading descriptions. After each evaluation, they publicly award (or withdraw) a 'badge' based on whether the distribution of reviews or complaints meets a set of *standards* decided by the platform. Finally, as discussed in the introduction, they offer some privacy: while participants observe all past complaints, these details are kept from customers. Motivated by these observations, we study the scope for a form of information design over $\underline{h}_t$ to improve all players' payoffs from long-run persuasion.

Consider a platform which can design the public record, $(r_t, \mathcal{R}_t)_t$. A **badge system** is a *binary* public history $(r_t, \mathcal{R}_t)_t$ where $\mathcal{R}_t = \{\mathcal{G}, \mathcal{B}\}$, $\forall t$. We interpret $\mathcal{G}$ as a public badge given to the sender (he is a 'badge bearer') and $\mathcal{B}$ as an absence thereof. A badge system is **simple** if it can be described by three parameters: *(i)* the length of **evaluation phases**, $\Gamma \in \mathbb{N}$; *(ii)* the length of **suspension phases**, $\beta \times \Gamma \in \mathbb{N}$; and *(iii)* a **set of standards**, $\mathcal{S}$, that reflect the outcomes required for a sender to retain a badge. More precisely the standard defines, for each possible $m \in \Delta\Theta$, a set $\mathcal{S}(m)$ of joint frequencies $s(m)$ over (m, ω).[19] As each message induces an action in equilibrium, we can interpret standards as targets for the reported experiences of receivers.

Given these parameters, a **simple badge system** works as follows: at $t = 1$, the sender receives the badge ($r_1 = \mathcal{G}$). This initiates an evaluation phase, throughout which the badge is retained. At the end of an evaluation phase, feedback obtained during that phase is compared against the standards. Specifically let

$$\ell_t(m, \omega) = \frac{1}{\Gamma} \sum_{\tau = t - \Gamma + 1}^{t} \mathbb{1}((m_t, \omega_t) = (m, \omega)) \tag{6}$$

be the realized frequency of (m, ω) in the Γ interactions preceding $t \geqslant \Gamma$, where $\mathbb{1}$ is the indicator function, and let $\ell_t(m) = (\ell_t(m, \omega))_{\omega \in \Omega}$. Then if $\ell_\Gamma(m) \notin \mathcal{S}(m)$ for some m, the sender loses the badge for $\beta \times \Gamma$ periods. Otherwise, his badge is renewed. After this, the entire process restarts.[20]

Our goal is to understand the scope of simple badge systems to improve persuasion. For this reason, we consider a setting in which arriving receivers do not initially know the date on which the sender is next evaluated. Similar to Kremer et al. (2014), we assume receivers arrive uniformly at random during evaluation phases and do not observe t. This simplifies analysis and only further limits receivers' information about $\underline{h}_t$, thereby leveraging the insight of section 3.1. In section 4, we show how platforms can create similar uncertainty if arrivals are nonrandom. We make two further assumptions.

Assumption 2 *The conditional distributions*

$$(p(\theta \mid \omega, a))_{\theta \in \Theta} \tag{7}$$

are linearly independent across Ω for all but at most one $a' \in A$. p is continuous in a.

Assumption 2 is a slight relaxation of the standard identifiability condition in repeated games.Without it the sender's strategy could not be accurately monitored even

[19]Each $s(m)$ is a feasible vector of joint probabilities over (m, ω), $\omega \in \Omega$, whose sum must of course not exceed 1. Formally each $s(m) \in \Delta_c \Omega := \{s \in \mathbb{R}^{|\Omega|} : \sum_{\omega \in \Omega} s(m, \omega) \leqslant 1, s \geqslant 0\}$.

[20]We use a system which immediately grants a badge and that has fixed length of suspension phases only to simplify the analysis. Our results do not hinge on these characteristics.

with access to infinite data. The relaxation allows for cases where some $a' \in A$ is uninformative about θ, such as Example 1. Second, we assume the following:

Assumption 3 *$V(\lambda)$ is continuous on an region whose closure contains λ for all λ.*

Assumption 3 is a mild payoff restriction. It rules out knife-edge cases where the receiver's action would radically change after *every* small perturbation of λ. But it still accommodates reasonable discontinuities in v (such as in Example 1) since it only requires continuity around *some* local perturbations of an information structure λ. As λ can be on the boundary of such regions, Assumption 3 is strictly weaker than continuity.

Let $\underline{\Lambda}(\mu_0)$ be the subset of $\Lambda(\mu_0)$ for which $V(\lambda)$ exceeds the sender's worst one-shot cheap talk payoff. Our second theorem establishes the value of simple badge systems:

Theorem 2 *Suppose Assumptions 2 and 3 hold. For any $\lambda^\star \in \underline{\Lambda}(\mu_0)$, there is a limiting equilibrium with a simple badge system in which, as $\delta \to 1$,*

1. *in an evaluation phase, a badge bearer's strategy implies a time-averaged information structure*

$$\lambda^\star = \frac{1}{\Gamma}\sum_{\tau=1}^{\Gamma} \mathbb{E}[\lambda_\tau|\mathcal{G}]$$

2. *average payoffs are $(U(\lambda^\star), V(\lambda^\star))$*

Theorem 2 shows that the benefits of coarse rating systems like those found online can be large: for patient senders, *any* commitment payoffs which improve on one-shot cheap talk are attainable. In stark contrast to section 2, simple badge systems allow the sender to capture the full rewards of his communication — indeed, he achieves the payoff $V(\lambda^\star)$ with a strategy that induces a time-averaged information structure of $\lambda^\star$ during evaluations. Moreover, it shows how badge systems can efficiently solve imperfect monitoring problems and support Pareto superior communication equilibria. For instance, in Example 1 the platform can sustain any payoffs in $\mathcal{CS}$ (see Figure 1b).

To prove the result, for any $\lambda^\star \in \underline{\Lambda}$ we construct a corresponding sequence of simple badge systems that admit good equilibria as $\delta \to 1$. To see the intuition, consider a system with long evaluation phases and standards which allow only distributions of

(m, ω) close to the one induced by $\lambda^\star$. Given Assumption 2, it is possible to choose standards tough enough that if the sender uses a time-average information structure that differs even a little from $\lambda^\star$ he will lose his badge, and yet lenient enough that if he adopts $\lambda^\star$ each period he will keep his badge with probability 1.[21]

Of course in *any* equilibrium, suspensions induce one-shot cheap talk incentives. To analyze a badge bearer's behavior we first adapt an argument from Radner (1985): If receivers' beliefs satisfy $m \approx \mu$, suspensions are long enough, and he is patient, then his optimal strategy is to play close to $\lambda^\star$ on average to avoid the punishment of a suspension. We then consider receivers' beliefs: if she sees a message m from a badge bearer, she must indeed infer it has been approximately generated by $\lambda^\star$. This is because she lacks any information about feedback history and arrival time. Hence, her beliefs do indeed satisfy $m \approx \mu$. Building on this intuition, the proof shows best responses and beliefs can be mutually closed in arbitrarily small intervals of each other. Along with Assumptions 2 and 3, we then prove existence via a fixed-point argument.

To see how these badge systems harness the insights of section 3.1, consider using a simple badge system to support $V(\lambda^\star) > \mu_0$ in Example 1. Trivially, if the seller's feedback has already met the standards on the last day of an evaluation phase then he will say "like new" irrespective of θ. This acts as a reward that the sender may get for good behavior earlier in the evaluation phase. The promise of this reward, and others like it, reduces the punishment necessary to incentivize credibility at earlier periods. Moreover, a single reward late on provides incentives at every prior period, allowing a reduction in punishment far greater than the actual reward. Consequently, on-path punishments can be pushed close to zero with relatively few rewards. Finally, by keeping customers in the dark, the badge system makes these rewards incentive compatible.

Interestingly, our analysis also implies the feasible set of equilibria is constrained by the *particular* choice of public record. Indeed, the platform controls the nature of communication in all equilibria where the sender seeks to retain the badge.

Remark 1 *Consider standards corresponding to equilibrium payoffs* $(U(\lambda^\star), V(\lambda^\star))$*. In any evaluation phase where a badge bearer's equilibrium strategy is not 'close to'* $\lambda^\star$ *on average, the average equilibrium payoffs must lie within the stage-Nash set.*

[21] Most platforms use standards that require 'good' feedback to exceed a threshold. In the context of Example 1, such standards can only sustain the efficient frontier of $\mathcal{CS}$.

Finally, we briefly discuss two aspects of our analysis. First, whether complete or incomplete, public record systems do require some source of commitment themselves. It is not our aim to explain that commitment but to tease out its implications for an uncommitted sender. Nonetheless, we note that the systems discussed are partitions of history. Thus, long-run incentives may serve as an efficient source of commitment. Moreover, the platform faces many copies of the same problem — hence, similar to Jackson and Sonnenschein (2007), the distribution of badges could be the source of discipline. Second, unlike the mediators typically analyzed in repeated games, our platform cannot commit to private messages. This restriction reflects most online platforms, who take a hands-off approach to individual transactions — perhaps because mediating every single interaction is too costly. By ignoring such mediation, we might worry that something lost. Yet despite limiting the power of the platform, Theorem 2 shows that badge systems can still implement the commitment set.

4 Discussion: Alternative Review Systems

In section 3.2 the platform's control over public records was strong: it could privately collect feedback from receivers, it could prevent receivers from accessing this feedback directly, and it benefited from the random arrival of receivers (further limiting their information). Moreover, communication was the only source of moral hazard. Here we discuss how review systems with less control can still improve outcomes. In doing so we also highlight the costs and benefits of each one relative to simple badge systems.

4.1 Deterministic Arrival Systems

In section 3.2 receivers could not infer the time until the next evaluation. While information on evaluation dates is often searchable, customer behavior online seems better approximated by uncertainty over such details. Our results support their apparent disinterest: eBay's strict standards incentivize badged sellers to almost always behave. Hence even small attention costs would deter a search. Yet with more relaxed standards, incentives to find out these dates could rise. We briefly discuss the implications of this here. The online Appendix provides formal results in the setting of Example 1.

As one might expect, such information can constrain the performance of some simple badge systems. Nonetheless, when t is observable the platform can replicate the effect of random arrivals with a system that instead randomizes evaluation dates directly. By hiding the (randomly chosen) evaluation dates from customers it is *as if* they had arrived randomly (see Proposition 6). Even if customers *must* be informed of evaluation dates, the platform can still replicate the required uncertainty by adopting more complex standards that depend on *full sequences* of outcomes rather than simple averages (see Proposition 7). The idea is to make the sender endogenously create the uncertainty by mixing between two non-stationary strategies.[22]

We view stochastic evaluation dates as just another interpretation of our random arrival assumption. However, simple badge systems have three advantages over more complex evaluations. First, they are easier to identify and implement. Second, they do not rely on common knowledge that the sender mixes with precisely the right probabilities. Finally, simple badge systems do not require exact knowledge of δ.

4.2 Blind Sender Review Systems

In many cases the platform's commitment to badge systems is easy to motivate. But in the absence of an ability to discipline a third-party platform, customers could instead maintain a distributed ledger of feedback themselves. Hence, we now consider whether payoffs can be improved when receivers observe all feedback but the sender only has limited access. We call this a **blind sender** review system. At each t the incoming customer observes $\underline{h}_t$. Badges are again awarded to the seller based on evaluations of standards, but now by the customers rather than a third party. The seller observes the badges provided but not the individual feedback, ω_t. Such systems *can* improve payoffs.

Proposition 5 *Consider Example 1 with imperfect monitoring, $p < 1$. Using a blind sender review system, full information payoffs (μ_0, μ_0) are attainable in the limit as $\delta \to 1$. The sender cannot earn payoffs greater than μ_0 in any Perfect Public Equilibrium.*

The proof uses a standard in which the seller loses his badge if and only if *every*

[22]Our construction is reminiscent of a mediated equilibrium proposed in Sugaya and Wolitzky (2018), with the additional constraint of sender indifference across strategies. We thank an anonymous referee for pointing us towards the latter alternative.

product sold receives bad feedback. Building on Abreu et al. (1991) and Fuchs (2007), we construct an equilibrium in which a badged seller is *continually* disciplined to be honest by "reusing" the *same* threat of punishment — dramatically reducing their aggregate cost. Moreover, there is now clearly no need to hide history from receivers.

Blind sender systems have two major limitations. As Proposition 5 shows, reusing punishments alone cannot costlessly support non-partitional information structures.[23] When customers observe a finer feedback history than the seller, the logic of Proposition 1 still applies. Moreover, they can be fragile: if a seller can find out that he has even a single good review, he can lie for the remainder of an evaluation without cost. This may be one reason why sellers do observe the history of feedback on most platforms.

4.3 Quota-Based Review Systems

Some platforms may receive very little feedback. Even without feedback, Margaria and Smolin (2018) show how to obtain a folk theorem in a setting with a long-run receiver and a sender with state-independent preferences. In the context of Example 1, their equilibrium constructions make the seller indifferent between all *sequences* of messages over long horizons. To generate indifference, the customer is required to occasionally buy regardless of the seller's claims. This is obviously not feasible with short-lived customers and complete public records. However, as the frequency of such 'rewards' is low as $\delta \to 1$, we conjecture a feedback-free badge system may be able to achieve any payoffs in $\mathcal{CS}$. Even if this is true, feedback-free ratings would still have limitations: they require exact knowledge of δ (to generate the indifference) and are not robust to state-dependent payoffs. By contrast simple badge systems are robust to both issues.

4.4 Simple Badge Systems for Other Forms of Moral Hazard

Online platforms face many forms of moral hazard beyond communication: sellers may fail to deliver on time, hosts may fail to keep accommodation clean, or workers may shirk. Simple badge systems can equally solve these problems. For the platform, the

[23] As Abreu et al. (1991) and Fuchs (2007) show, all gains from reusable punishments can be realized within a Perfect Public Equilibrium. Any gains that can be realized by allowing the seller to condition on his private history would necessarily come from the reward mechanism described in section 3.1.

key difference between communication and these classic moral hazards lies in how much players can infer beyond the badge. In our setting additional historical information may leak out (receivers can update after seeing the sender's message). But classic moral hazards do not suffer this leakage, so the argument underpinning Theorem 2 applies to them *a fortiori* (see Corollary 5 in the online Appendix). Indeed, subject to adjustments of Assumptions 2 and 3, simple badge systems ought to support the commitment set for any two-player stage game in the environment of Fudenberg and Levine (1994).

5 Related Literature

Since Kamenica and Gentzkow (2011), a large body of work has sought to understand strategic communication when a sender has exogenous commitment power.[24] We contribute to this literature by relaxing the commitment assumption. Perez-Richet (2014), Salamanca (2019), Lipnowski et al. (2019) and Perez-Richet and Skreta (2018) study persuasion when the sender has partial commitment power. Mathevet et al. (2019) consider how well repetition substitutes for commitment when the sender may be an exogenous 'commitment type': if the probability of a commitment type is not too low, the sender obtains the Stackelberg payoff. We examine persuasion when there is no hope of the sender having exogenous commitment power. Moreover, we show that the design of coarse public rating systems (and repetition) can substitute for a complete lack of commitment on the sender's part. Like us, Chakraborty and Harbaugh (2010) and Lipnowski and Ravid (2019) analyze persuasion when the sender cannot commit. Relative to them, our contribution is to introduce an endogenous source of commitment.

Partitional information structures have gained attention for their simplicity, applicability and tractability. Kolotilin (2018), Dworczak and Martini (2019), Mensch (2019) and Kolotilin and Li (2019) identify conditions under which *monotone* partitional information structures are optimal.[25] We identify a new, enforcement-based appeal of partitional information structures and provide new conditions for their optimality.

We also relate to recent work on cheap talk in repeated games. Hörner et al. (2015)

[24]For instance, Rayo and Segal (2010), Taneva (2019), Bergemann and Morris (2013), Ely (2017), and Bizzotto et al. (2019).

[25]By contrast, non-partitional information appears in the Stackelberg solutions to Example 1, Brocas and Carrillo (2007), Gill and Sgroi (2008, 2012), Rayo and Segal (2010) and Alonso and Câmara (2016).

and Margaria and Smolin (2018) examine such games between long-run players, where a folk theorem obtains. Like us, Jullien and Park (2019) examine a setting with short-run receivers. These papers all consider settings where truthful equilibria are without loss. In our setting there is a meaningful role for persuasion. Aumann and Maschler (1995) study a repeated zero-sum game with infinitely-patient agents and asymmetric information about *perfectly persistent* stage payoffs. In such settings, they show that Stackelberg persuasion is a credible cheap-talk outcome. Hence, there is no need for future credibility concerns to sustain persuasion. We focus on the complementary settings where repetition is necessary, and where $\delta < 1$.[26] Kuvalekar et al. (2019) study a problem similar to ours, but with no monitoring at all of the sender's accuracy.

Other papers have examined hiding history in repeated games. Sugaya and Wolitzky (2018) shows how private monitoring with a mediator can help sustain a collusive agreement ('market segmentation') by making it more difficult for firms to tailor deviations to market conditions.[27] To isolate this channel, they assume perfect monitoring. By contrast, we show how hiding history can eliminate the inefficiencies caused by imperfect monitoring. Abreu et al. (1991) and Fuchs (2007) hide history from *long-run* players when monitoring is imperfect. They show inefficiencies can be reduced by "reusing punishments". Our badge systems cannot reuse punishments as the sender observes the full history. Instead, we show how a platform can persuade receivers to provide rewards rather than punishments. This allows us to tackle *both* the imperfect monitoring and mixing problems present in Fudenberg and Levine (1994). Bhaskar and Thomas (2019) study a repeated investment game with bounded memory. They show coarsening records can sustain investment by making punishments sequentially rational. While this problem could also be solved by providing the full history, in our model coarser ratings strictly improve outcomes. Furthermore, payoffs in their setting remain bounded by a logic akin to Fudenberg et al. (1990). Finally, we contribute to this literature by showing how our methods can characterize the full limiting set of equilibrium payoffs.

Several papers, going back to Myerson (1986), have designed information in extensive-form games to improve payoffs. Recently, Gershkov and Szentes (2009) apply this to voting mechanisms and Kremer et al. (2014) to social learning. We show how this idea

[26]If our sender were *infinitely* patient, he could also secure Stackelberg payoffs with complete records. This follows from a quota-based argument rather than Aumann and Maschler (1995)'s insight.

[27]While their main results apply to a stochastic game, they do provide a numerical example where mediation can improve payoffs in a stationary environment. In this example, monitoring is perfect.

can be used to improve persuasion and the provision of efficient incentives in repeated games more generally. Doval and Ely (2020) provide related tools for finite games, but these do not apply to our infinite-horizon setting.

Dellarocas (2005) and Tadelis (2016) examine how online review systems can implement the Fudenberg and Levine (1994) bound. Our results show how review systems can do better. Ekmekci (2011) shows how reducing a rating system's memory can prevent the collapse of reputations, à la Cripps et al. (2004), and improve seller payoffs; but the existence of reputational types is necessary for making improvements. By contrast, our badge systems yield Pareto improvements without any types at all.

6 Conclusion

We have analyzed a model with a meaningful role for persuasion where commitment lies in repetition and a public record of past accuracy. When the public record is complete, we gave necessary and sufficient conditions on the underlying persuasion problem under which a loss of ex ante commitment is costly. Unfortunately, even if the sender's private information can be observed ex post, these conditions are often not met. However, we show that by treating the public record as a complementary information design problem Pareto gains are possible. Indeed, a platform can almost perfectly substitute for the sender's lack of commitment via badge systems similar to those online.

We conclude with some open questions. First, instead of examining cheap talk, we could have examined long-run persuasion when the sender's private information is endogenously acquired and/or partially verifiable. It is easy to construct examples where such changes significantly alter communication outcomes to the detriment of both parties. Second, how can commitment to the rules that generate the public record be justified? As is common in the literature, this paper treated it as a given but we did suggest two possible avenues: the platform as a third long-run player and the distributed ledger interpretation of section 4.2. Exploring this second possibility may be of particular value in the context of block-chain technologies. Finally, Remark 1 notes how a simple badge system restricts the set of feasible equilibrium outcomes. This raises the intriguing question of exactly how much power a platform has to implement their preferred equilibrium outcomes via their control over the public record.

References

Dilip Abreu, Paul Milgrom, and David Pearce. Information and timing in repeated partnerships. *Econometrica*, 59(6):1713–1733, 1991.

Ricardo Alonso and Odilon Câmara. Persuading voters. *The American Economic Review*, 106(11):3590–3605, 2016.

Robert J Aumann and Michael B Maschler. *Repeated games with incomplete information*. MIT press, 1995.

Dirk Bergemann and Stephen Morris. Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308, 2013.

James Best and Daniel Quigley. Persuasion for the long run. *Working Paper*, 2017.

V Bhaskar and Caroline Thomas. Community enforcement of trust with bounded memory. *The Review of Economic Studies*, 86(3):1010–1032, 2019.

Jacopo Bizzotto, Jesper Rüdiger, and Adrien Vigier. Dynamic persuasion with outside information. *Working Paper, Available at SSRN 3154313*, 2019.

Isabelle Brocas and Juan D Carrillo. Influence through ignorance. *RAND Journal of Economics*, 38(4):931–947, 2007.

Archishman Chakraborty and Rick Harbaugh. Persuasion by cheap talk. *American Economic Review*, 100(5):2361–2382, 2010.

Martin W Cripps, George J Mailath, and Larry Samuelson. Imperfect monitoring and impermanent reputations. *Econometrica*, 72(2):407–432, 2004.

Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38(10):886–893, 1952.

Chrysanthos Dellarocas. Reputation mechanism design in online trading environments with pure moral hazard. *Information Systems Research*, 16(2):209–230, 2005.

L Doval and J Ely. Sequential information design. *Working Paper*, 2020.

Piotr Dworczak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy*, 127(5):1993–2048, 2019.

Mehmet Ekmekci. Sustainable reputations with rating systems. *Journal of Economic Theory*, 146(2):479–503, 2011.

Jeffrey C Ely. Beeps. *The American Economic Review*, 107(1):31–53, 2017.

Ky Fan. Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences of the United States of America*, 38 (2):121, 1952.

William Fuchs. Contracting with repeated moral hazard and private evaluations. *American Economic Review*, 97(4):1432–1448, 2007.

D Fudenberg and D Levine. Efficiency and observability with long run and short run players. *Journal of Economic Theory*, 62:103–135, 1994.

Drew Fudenberg, David M Kreps, and Eric S Maskin. Repeated games with long-run and short-run players. *The Review of Economic Studies*, 57(4):555–573, 1990.

Matthew Gentzkow and Emir Kamenica. Costly persuasion. *American Economic Review*, 104(5):457–462, 2014.

Alex Gershkov and Balázs Szentes. Optimal voting schemes with costly information acquisition. *Journal of Economic Theory*, 144(1):36–68, 2009.

David Gill and Daniel Sgroi. Sequential decisions with tests. *Games and Economic Behavior*, 63(2):663–678, 2008.

David Gill and Daniel Sgroi. The optimal choice of pre-launch reviewer. *Journal of Economic Theory*, 147(3):1247–1260, 2012.

Irving L. Glicksberg. A further generalization of the kakutani fixed point theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.

Johannes Hörner, Satoru Takahashi, and Nicolas Vieille. Truthful equilibria in dynamic bayesian games. *Econometrica*, 83(5):1795–1848, 2015.

Xiang Hui, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan. Reputation and regulations: evidence from ebay. *Management Science*, 62(12):3604–3616, 2016.

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

Bruno Jullien and In-Uck Park. Communication, feedbacks and repeated moral hazard with short-lived buyers. *Working Paper*, 2019.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

Anton Kolotilin. Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635, 2018.

Anton Kolotilin and Hongyi Li. Relational communication. *Working Paper*, 2019.

Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the Wisdom of the Crowd. *Journal of Political Economy*, 122(5):988–1012, 2014.

Aditya Kuvalekar, Elliot Lipnowski, and Joao Ramos. Goodwill in communication. *Working Paper, Available at SSRN 3435196*, 2019.

Elliot Lipnowski and Doron Ravid. Cheap talk with transparent motives. *Working Paper, Available at SSRN 2941601*, 2019.

Elliot Lipnowski, Doron Ravid, and Denis Shishkin. Persuasion via weak institutions. *Working Paper, Available at SSRN 3168103*, 2019.

Chiara Margaria and Alex Smolin. Dynamic communication with biased senders. *Games and Economic Behavior*, 110:330–339, 2018.

Laurent Mathevet, David Pearce, and Ennio Stacchetti. Reputation and information design. *Working Paper*, 2019.

Jeffrey Mensch. Monotone Persuasion. *Working Paper*, 2019.

Roger B Myerson. Multistage games with communication. *Econometrica*, 52(2):323–358, 1986.

Chris Nosko and Steven Tadelis. The limits of reputation in platform markets: An empirical analysis and field experiment. Technical report, National Bureau of Economic Research, 2015.

Eduardo Perez-Richet. Interim Bayesian persuasion: First steps. *American Economic Review*, 104(5):469–474, 2014.

Eduardo Perez-Richet and Vasiliki Skreta. Test design under falsification. *Working Paper*, 2018.

Roy Radner. Repeated Principal-Agent Games with Discounting. *Econometrica*, 53(5): 1173–1198, 1985.

David Rahman. The power of communication. *American Economic Review*, 104(11): 3737–51, 2014.

Luis Rayo and Ilya Segal. Optimal information disclosure. *Journal of Political Economy*, 118(5):949–987, 2010.

Andrés Salamanca. The value of mediated communication. Working Paper, 2019.

Takuo Sugaya and Alexander Wolitzky. Maintaining privacy in cartels. *Journal of Political Economy*, 126(6):2569–2607, 2018.

Steven Tadelis. Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340, 2016.

Ina Taneva. Information design. *American Economic Journal: Microeconomics*, 11(4): 151–85, November 2019.

Appendix: Proofs of Main Results

Proposition 2

By definition, $C(\lambda) \geqslant 0$ for all $\lambda \in \Lambda(\mu_0)$. Thus $V^{LP}(\mu_0) = \hat{v}(\mu_0)$ if and only if $C(\hat{\lambda}) = 0$. Hence, either (i) $\tilde{v}(\mu_0) = \hat{v}(\mu_0)$, where $\tilde{v}$ is as defined in (3), or (ii) there exists a pair $\mu,\ \mu' \in \operatorname{supp} \hat{\lambda}$ such that $v(\mu) = v(\mu')$. But (ii) violates Assumption 1.[28] $\square$

[28]For instance, consider $\mu_0 = \frac{1}{2}\mu + \frac{1}{2}\mu'$.

Theorem 1 and Corollary 1

Before establishing the argument, we first make two simple observations:

Lemma 2 *Consider two functions, f, g on some set S. If $f \geqslant g$ everywhere, then cav $f \geqslant$ cav g.*

Lemma 2 is a well-known property of concave envelopes. Hence we omit the proof.

Lemma 3 $v \leqslant \tilde{v} \leqslant V^{LP} \leqslant \hat{v}$ *everywhere on* $\Delta\Theta$.
Proof: The first inequality follows because no information is a partitional information structure. The second follows from *(i)* (LP), *(ii)* $\tilde{\Lambda}(\mu_0) \subset \Lambda(\mu_0)$ and *(iii)* $C(\lambda) = 0$ for $\lambda \in \tilde{\Lambda}$, and the third from (LP). $\square$

Notice that Lemmas 2 and 3 imply *cav* $\tilde{v} =$ *cav* $V^{LP} = \hat{v}$. We now prove both Theorem 1 and Corollary 1. Suppose that $\tilde{v}$ is concave. Then $\tilde{v} = \text{cav } \tilde{v} = \hat{v}$. From Lemma 3, $V^{LP} = \hat{v}$ everywhere and hence V^{LP} is concave, too. For the converse, suppose that $\tilde{v}$ is not concave. Then there exists at nonempty set $A \subset \Delta\Theta$ such that $\tilde{v}(\mu) < \text{cav } \tilde{v}(\mu) = \hat{v}(\mu)$ if and only if $\mu \in A$. Thus, $\hat{\lambda}$ is non-partitional iff $\mu_0 \in A$. Hence by Proposition 2, $V^{LP}(\mu) < \hat{v}(\mu)$ iff $\mu \in A$. Thus, V^{LP} is not everywhere equal to *cav* $\tilde{v}$ and hence, is not concave.

Finally to complete Corollary 1, note from Lemma 3 that $\hat{v} - V^{LP} \leqslant \text{cav } \tilde{v} - \tilde{v}$. $\square$

Corollary 3

Without loss, assume that each $a \in A$ is chosen be the receiver for some $\mu \in \Delta\Theta$. Let $a^\star$ be the sender-preferred action and define $M(a^\star) := \{\mu : a^\star \in \arg\max_{a \in A} \sum_{\theta\in\Theta} \mu(\theta) u(a', \theta)\}$ as the (nonempty) subset of $\Delta\Theta$ on which the receiver chooses $a^\star$. Note that $M(a^\star)$ is compact and convex. By Assumption 1, either (i) $a^\star$ is uniquely sender-optimal or (ii) the sender is indifferent between all actions. We focus on case (i) and show that $V^{LP} \neq \tilde{v}$. Generically, $M(a^\star)$ contains an open subset. Therefore, there exists some belief $\mu' \in \text{int } \Delta\Theta$ on the boundary of $M(a^\star)$ at which the receiver is indifferent between $a^\star$ and some $a' \neq a^\star$. Now consider a sequence $\{\mu_n\}_n$ such that $\mu_n \notin M(a^\star)$, $\forall n$, and $\mu_n \to \mu'$. For each μ_n, we have by definition

$$\tilde{v}(\mu_n) = \max_{\mathcal{P} \in B_\Theta} \sum_{i=1}^{k} \mu(P_i) v(\mu_{P_i}) = v(a^\star) - \sum_{i=1}^{k} \mu(P_i)(v(a^\star) - v(\mu_{P_i}))$$

where $\mu(P_i) := \sum_{\theta\in P_i} \mu(\theta)$, and μ_{P_i} is the projection of μ_0 into ΔP_i. Now the convexity of $M(a^\star)$ and $\mu_n \notin M(a^\star)$ imply that for *any* partition $\mathcal{P}$ there exists at least one posterior $\mu_{P_i} \notin M(a^\star)$. But since the sender is not indifferent across all actions, there must be a $d > 0$ such that $v(a^\star) - v(a) \geqslant d$ for all $a \in A \backslash \{a^\star\}$. Hence for any $\varepsilon > 0$ and all n sufficiently large we must have $v_{\mathcal{P}} \leqslant v(a^\star) - d\,(\cdot \min_\theta \mu'(\theta) - \varepsilon)$. But μ' is interior. Hence $\lim \tilde{v}(\mu_n) < v(\mu')$. Since $\tilde{v}$ is not continuous, it cannot be convex. The corollary then follows from Theorem 1. $\square$

Proposition 3

Note that $\lambda_\mu^{\mathcal{P}}$ is just a lottery over k posteriors, $\{\mu_1, \ldots, \mu_k\}$, where $\mu_j(\theta) = \dfrac{\mu(\theta)}{\sum_{\theta' \in P_j} \mu(\theta')}$ if $\theta \in P_j$ and 0 otherwise. By Bayes plausibility, $\lambda_j := \lambda_\mu^{\mathcal{P}}(\mu_j) = \sum_{\theta \in P_j} \mu(\theta)$. Let $v_{\mathcal{P}}(\mu) := \sum_{j=1}^k \lambda_j v(\mu_j)$ be the sender's payoff from adopting $\mathcal{P}$.

(*If*) Given (i), we must have $v_{\mathcal{P}} \geqslant v$ everywhere. And by definition, $\tilde{v} \geqslant v_{\mathcal{P}}$. Thus from Lemmas 2 and 3 we have cav $v_{\mathcal{P}} = \hat{v}$. Thus, if we can show $v_{\mathcal{P}}$ is concave we are done (by Theorem 1). Consider priors μ_0, μ_0' and $\mu_0'' = \alpha\mu_0 + (1-\alpha)\mu_0'$ and let $\gamma_j = \dfrac{\alpha\lambda_j}{\alpha\lambda_j + (1-\alpha)\lambda_j'}$, where λ' corresponds to μ_0'. Then

$$
\begin{aligned}
v_{\mathcal{P}}(\mu_0'') &= \textstyle\sum_j \lambda_j'' v(\mu_j'') \\
&= \textstyle\sum_j \Big(\alpha\lambda_j + (1-\alpha)\lambda_j'\Big) v(\mu_j'') \\
&\geqslant \textstyle\sum_j \Big(\alpha\lambda_j + (1-\alpha)\lambda_j'\Big) \Big(\gamma_j v(\mu_j) + (1-\gamma_j) v(\mu_j')\Big) \\
&= \textstyle\sum_j \Big(\alpha\lambda_j v(\mu_j) + (1-\alpha)\lambda_j' v(\mu_j')\Big) \\
&= \alpha v_{\mathcal{P}}(\mu_0) + (1-\alpha) v_{\mathcal{P}}(\mu_0'),
\end{aligned}
$$

where the second line follows from $\lambda_j'' = \alpha\lambda_j + (1-\alpha)\lambda_j''$, and the inequality from $\mu_j'' = \gamma_j\mu_j + (1-\gamma_j)\mu_j'$ and (ii). $\square$

(*Only if*) Suppose (i) is violated. Then there exists some μ_0 such that $v(\mu_0) > v_{\mathcal{P}}(\mu_0)$. Therefore, $\hat{v}(\mu_0) > v_{\mathcal{P}}(\mu_0)$ too, since $\hat{v} \geqslant v$ — contradicting optimality of $\mathcal{P}$. Similarly if (ii) is violated then there exists some j and $\mu_0 \in \Delta P_j$ such that $\hat{v}(\mu_0) > v(\mu_0) = v_{\mathcal{P}}(\mu_0)$ since by definition v and $v_{\mathcal{P}}$ coincide on any ΔP_j, $j = 1, \ldots, k$. $\square$

Proposition 4

First we define two (pure) communication strategies at the stage game. Let σ^T be the *truthful* strategy, where σ^T('like new' $\mid \theta) = 1$ if $\theta = h$ and 0 otherwise, and let σ^b be the babbling strategy σ^b('like new' $\mid \theta) = 1$ for $\theta = h, l$. We construct an equilibrium of the incomplete record game where the sender uses either σ^T or σ^b as a function of the public record, r_t. For brevity, we consider strategies which from $t = 3$ onward depend on the public record indirectly via a public state variable, $s_t \in \{s^T, s^b\}$, which evolves as follows: at $t = 3$, if $r_3 = A$ then $s_t = s^T$. If $r_3 = B$ then $s_3 = s^T$ with probability $1 - q \geqslant 0$; and if $r_3 = C$ $s_3 = s^T$ with probability $1 - q - r \geqslant 0$ for parameters $q, r \geqslant 0$. At $t > 3$, if $s_{t-1} = s^b$ then $s_t = s^b$. However, if $s_{t-1} = s^T$ then $s_t = s^T$ with certainty if $\omega_t \neq b$ and with probability $1 - r$ if $\omega_t = b$.[29]

Consider the following strategy profile. The sender plays σ^T if $t = 1$, if $t = 2$ and $\omega_1 = b$, and if $s_t = s^T$, $t \geqslant 3$. Otherwise, he plays σ^b. If $t \in \{1, 2\}$ or $s_t = s^T$, then receiver buys if and only if m_t = 'like new'. Otherwise she never buys, regardless of m_t.

[29] Note we implicitly use a public randomization device here.

Finally set $q = (1-\delta)\frac{1-\delta(p+\mu_0-2p\mu_0)}{\delta^2\mu_0(2p-1)}$, $r = \frac{1-\delta}{\delta\mu_0(2p-1)}$ and $l = -1$. Note $q, r \geqslant 0$, since $p + \mu_0 - 2\mu_0 p = p(1\mu_0) + (1-p)\mu_0$ and $p > \frac{1}{2}$. Moreover for $p > \frac{1}{2}$, $\mu_0 \in (0,1)$, q, r are continuous in δ, with $q = r = 0$ at $\delta = 1$. Thus there is a threshold $\underline{\delta}(p, \mu_0) < 1$ s.t. $q, r \in (0,1)$ for $\delta \geqslant \underline{\delta}(p, \mu_0)$.

We first verify equilibrium. Trivially each receiver at $t \neq 2$ is playing a best response. It is also easy to show the sender's strategy is a best response for $\delta \geqslant \underline{\delta}(p, r)$.[30] At $t = 2$, the receiver's best response is indeed to buy iff $m_2 =$ 'like new', so long as:

$$\Pr[\theta = h \mid m_2 = \text{`like new'}] = \frac{\mu_0}{1-(1-\mu_0)\Pr[\omega_1 = b]} = \frac{\mu_0}{1-(1-\mu_0)\mu_0(1-p)} \geqslant \frac{1}{2}. \quad (8)$$

For any $p > \frac{1}{2}$, the LHS is continuous in μ_0 and (8) holds strictly at $\mu_0 = \frac{1}{2}$. Thus for μ_0 close enough to $\frac{1}{2}$ the second receiver best responds too.

Let $V^T = \frac{1}{1-\delta}\mu_0(1 - \frac{1-p}{p})$. Applying Fudenberg and Levine (1994), it is easy to see V^T is the sender's maximal payoff in any equilibrium with a complete public record. We now show players can be better off on average. The sender's payoff is:

$$V' = V^T + \delta\left(1 - \mu_0 + \mu_0\frac{1-p}{p}\right) > V^T. \quad (9)$$

The discounted sum of receivers' utility is:[31]

$$U' = U^T + \delta\left(\mu_0\frac{1-p}{p} - \left(\Pr[\omega_1 \neq b] - \Pr[\omega_1 = b]\right)(1-\mu_0)|l|\right) \quad (10)$$

where $\Pr[\omega_1 = b] = \mu_0(1-p)$, $U^T = \frac{1}{1-\delta}\mu_0(1 - \frac{1-p}{p})$. On rearrangement, $U' > U^T$ iff

$$\frac{1-p}{p} > (1 - 2\mu_0(1-p))\frac{1-\mu_0}{\mu_0}. \quad (11)$$

Clearly both the LHS and RHS are continuous in p, μ_0, for $\mu_0 > 0$. At $\mu_0 = \frac{1}{2}$, (11) is satisfied strictly for $\frac{1}{2} \leqslant p < \frac{\sqrt{5}-1}{2}$. Thus clearly there exist μ_0, p close enough to $\frac{1}{2}$ such that both (8) and (11) hold. Indeed both are satisfied for $p = \frac{6}{11}$, $\mu_0 = \frac{5}{11}$. Finally, for any such choice this is an equilibrium so long as $1 > \delta \geqslant \underline{\delta}(p, \mu_0)$. For $p = \frac{6}{11}$, $\mu_0 = \frac{5}{11}$, this becomes $\delta \geqslant \frac{121}{126}$). $\square$

Theorem 2

Fix some $\lambda^\star \in \underline{\Lambda}(\mu_0)$ and write the corresponding payoffs $u^\star = U(\lambda^\star)$, $v^\star = V(\lambda^\star)$. Fixing $\varepsilon > 0$, we show that there is a threshold $\delta_\varepsilon < 1$ such that when $\delta \geqslant \delta_\varepsilon$ a simple badge system (SBS) and a corresponding equilibrium exists in which the discounted average payoffs (u, v) satisfy $||(u,v) - (u^\star, v^\star)|| \leqslant \varepsilon$. Hence taking limits as $\varepsilon \to 0$ establishes the claim. First, we set up some preliminary concepts.

[30] Available on request.

[31] We discount future receivers for illustrative simplicity: Theorem 2 holds even if one takes limiting time-average receiver payoffs.

Preliminary: r-recursive strategy

An r-recursive strategy for the sender depends only on the current phase r_t and the outcomes within that phase. Specifically, an r-recursive strategy defines functions:

$$\sigma^r : \bigcup_{l=0}^{\Gamma_r - 1} \mathcal{H}_l \rightarrow \Delta\Delta\Theta,\ r = \mathcal{G}, \mathcal{B}. \tag{12}$$

Let h'_l be the sub-component of h_t formed from the most recent $l \leqslant t$ periods. In a r-recursive strategy, the sender's behavior $\sigma\left(r_t, h'_l\right)$ depends on the current badge status, r, and also on previous outcomes within the current phase. If the current badge was first awarded l periods ago, this is just h'_l, $l = 0, \ldots, \Gamma_r - 1$.

Given r and message m, a receiver forms her belief $\mu^r(m)$. Given the SBS, this may only depend on r_t and m in equilibrium. Hence, the sender always has a best response which is r-recursive too. Specifically, we will prove our result holds within the class of r-recursive trigger strategies: players coordinate on the sender-worst stage Nash equilibrium whenever $r_t = \mathcal{B}$. Each such period players earn stage payoffs $\left(u^{NE}, v^{NE}\right)$, respectively, which for ease we normalize to $(0, 0)$.

If players adopt r-recursive trigger strategies, the senders' discounted average payoffs $\mathcal{V}$ are simply:

$$\mathcal{V}\left(\sigma, \mu^{\mathcal{G}}\right) = \sum_{t=0}^{\Gamma_{\mathcal{G}-1}} \delta^t \mathbb{E}_{\sigma_t^{\mathcal{G}}} v\left(\mu^r\left(m_t\right)\right) + \delta^\Gamma \left(\Pr_\sigma\left[r_\Gamma = \mathcal{G}\right] \mathcal{V}\left(\sigma, \mu^{\mathcal{G}}\right) + \delta^{\beta\Gamma} \Pr_\sigma\left[r_\Gamma = \mathcal{B}\right] \mathcal{V}\left(\sigma, \mu^{\mathcal{G}}\right)\right) \tag{13}$$

Preliminary: λ-Tight Standards

Let d_n be the Euclidean metric on $\mathbb{R}^n$. We consider a class of SBS which use the following standards:

$$\mathcal{S} = \left\{ h'_\Gamma : d_1\left(\frac{\sum_{t=0}^{\Gamma-1} \mathbb{1}(m, \omega_j)}{\Gamma}, q\left(m, \omega\right)\right) \leqslant \begin{cases} \chi, & \text{if } \mu = m \in \operatorname{supp}(\lambda), \\ 0, & \text{otherwise.} \end{cases} \right\} \tag{14}$$

for some choice $\lambda \in \Lambda\left(\mu_0\right)$, $\chi > 0$, where

$$q\left(m, \omega\right) := \sum_\Theta p\left(\omega, \theta \mid a\left(\mu\right)\right) \lambda(m \mid \theta)$$

is the joint probability of observing m and ω implied by information structure λ.

Proof.

We are now ready to prove the result. Unlike repeated games with two long-run players, we cannot use threats of large continuation punishments to make receivers behave as we wish. To establish existence of an equilibrium with the desired properties, we cannot lean on standard fixed point arguments either – indeed, a 'bad' equilibrium (babbling) always exists. Finally, we cannot leverage contraction mapping arguments (indeed, our result proves the sender's value function cannot be a contraction). We overcome these difficulties in three steps, showing that players' best responses can be trapped in a mutually closed subspace which can be made arbitrarily small about $\lambda^\star$ as $\delta \to 1$. The result then follows from existence arguments applied to this restricted domain.

Verifying equilibrium in $\mathcal{B}$-phases is trivial and thus omitted. We focus on $\mathcal{G}$-phases from now, and so drop 'r' superscripts where the meaning is clear. Given the sender uses r-recursive strategy σ, a receiver's beliefs on observing on-path message m are:[32]

$$\mu(\theta \mid m; \sigma) = \Pr[\theta \mid m] = \frac{\sum_{l=0}^{\Gamma-1} \Pr_\sigma[\theta, m, l]}{\sum_{\Theta} \sum_{l=0}^{\Gamma-1} \Pr_\sigma[\theta, m, l]}. \tag{15}$$

For δ large, our aim is to establish existence of an SBS with a corresponding equilibrium in which the sender's average information structure is approximately $\lambda^\star$. Given constants $\varepsilon_S, \varepsilon_R > 0$, we are interested in strategies and beliefs which satisfy the following bounds (for all on-path m):

$$d_1\left(\mathbb{E}\left[\frac{\sum_{l=0}^{\Gamma} \sigma_l(m \mid h_l', \theta)}{\Gamma} \;\Big|\; \theta\right], \lambda^\star(m \mid \theta)\right) \leqslant \varepsilon_S, \quad d_N(\mu^r(m) - m) \leqslant \varepsilon_R, \tag{16}$$

where $N = |\Theta|$ and $\sigma_l(m \mid h, \theta)$ is the probability of sending m given h, θ.

Notice that the set of σ and μ functions satisfying (16) is compact and convex. From Assumption 3, for small $\varepsilon > 0$, we can also always find a $\lambda' \in \Lambda(\mu_0)$ such that (i) $||V(\lambda') - V(\lambda^\star)|| \leqslant \frac{\varepsilon}{2}$; and (ii) $V(\lambda)$ is continuous on a closed $\frac{\varepsilon}{2}$-ball around λ'. Without loss for payoffs, take λ' to have a finite support.[33] Given this, Assumption 3 moreover implies there exists a finite $\bar{\epsilon}(\lambda') > 0$ such that V is continuous for $|\lambda - \lambda'| \leqslant \bar{\epsilon}(\lambda')$, and v is continuous for $|\mu^r(m) - m|$, for all $m \in \text{supp } \lambda'$.

Given this, (13) is clearly continuous in (σ, μ) around λ' and quasi-concave in σ, while (15) is continuous in σ. Given any $0 < \varepsilon_S, \varepsilon_R \leqslant \bar{\epsilon}(\lambda')$, we show in three steps that for δ sufficiently high and $\lambda' \in \underline{\Lambda}(\mu_0)$ sufficiently close to $\lambda^\star$, we can choose λ'-tight standards such that the sender's best response and (15) can be mutually bounded by (16). Step 1 establishes receiver beliefs can be closed in this way. Given this, step 2 establishes a patient sender wants his *feedback* to meet the standards with high probability. Step 3 proves meeting the standards requires he uses a *strategy* that is indeed closed by (16). We conclude that an equilibrium exists with strategies and beliefs in this region (Debreu (1952), Glicksberg (1952), Fan (1952)).[34] The rest of the proof then follows simply.

[32]Off-path m will be treated as a random draw from the sender's worst Nash information structure. This, and the standards, ensure such a deviation is never profitable.

[33]The standard Carathéodory argument (Kamenica and Gentzkow, 2011) applies here.

[34]This establishes Nash existence. Here, a Nash equilibrium is also a PBE because (i) receivers cannot observe the sender's previous actions within a $\mathcal{G}$-phase, (ii) off-path deviations are never profitable.

1. Verify bounds on receivers' best responses

Given the sender's strategy σ, define the time-average strategy during a $\mathcal{G}$-phase as

$$\bar{\sigma}(m,\theta) := \frac{\sum_{l=0}^{\Gamma} \sigma_l^{\mathcal{G}}(m \mid \theta)}{\Gamma}. \tag{17}$$

We first show that, if $\mathbb{E}[\bar{\sigma} \mid \mathcal{G}]$ is bounded within an ε_S-ball of λ', for ε_S sufficiently small, then the receiver's belief $\mu(m)$ can indeed be bounded to stay within ε_R of m, for all $m \in \operatorname{supp}\{\lambda'\}$. This follows from (15). For each θ, the numerator of this expression can be written $\mu_0(\theta) \sum_{l=0}^{\Gamma-1} \mathbb{E}_\sigma[\bar{\sigma}(m,\theta)]$, which approaches $\mu_0(\theta)\lambda'(m \mid \theta)$ as $\tilde{\varepsilon}_S \to 0$. Since $\sum_\Theta \mu_0(\theta)\lambda'(m \mid \theta) > 0$ for all $m \in \operatorname{supp}\{\lambda'\}$, (15) thus converges to m.

2. Bounding sender's incentive to deviate from standards

Fix λ' such that $|V(\lambda') - V(\lambda^\star)| \leqslant \frac{\varepsilon}{2}$ and about which V is locally continuous. Choose $0 < \varepsilon_R < \bar{\epsilon}(\lambda')$ small enough that sender's stage payoff from using λ' exceeds his worst stage Nash payoff, for any receiver beliefs bounded from m by ε_R:[35]

$$\min_{\mu(m):d(\mu,m)\leqslant\varepsilon_R} \mathbb{E}_{\lambda'}[v(\mu^r(m))] > 0.$$

Now consider the class of λ'-tight standards. By a straightforward extension of Radner (1985), for any constant $z > 0$ we can find $\Gamma(z)$, $\chi(z) > 0$ and corresponding $\delta_\Gamma < 1$, $\beta_\delta > 0$, such that if $\Gamma \geqslant \Gamma(z)$, $\delta \geqslant \delta_\Gamma$ the sender's optimal strategy, σ, satisfies

$$\Pr_\sigma\left[h'_\Gamma \in \mathcal{S}_z \mid r = \mathcal{G}\right] \geqslant 1 - z$$

in a $\mathcal{G}$-phase, where $\mathcal{S}_z$ sets $\Gamma = \Gamma(z)$, $\chi = \chi(z)$ in (14). The proof of this fact is a simple extension of Radner (1985), and hence omitted.

3. Bounding the sender's average strategy in a $\mathcal{G}$-phase

We now show that, for any ε_S there is some $z > 0$ sufficiently small and some $\Gamma \geqslant \Gamma(z)$, such that a sender meets the standard with probability exceeding $1 - z$ if and only if his strategy satisfies (16). To do that, we first argue that as $z \to 0$ strategies σ_z meet $\mathcal{S}_z$ with probability exceeding $1 - z$ if and only if:

$$d_1(\mathbb{E}_{\sigma_z}[\ell], q) \leqslant \eta_z,$$

for a bound $\eta_z \to 0$ as $z \to 0$ (i.e., he will be close to the standards *in expectation*).

(If) The sender can always choose to play $\sigma_l(h'_l) = \lambda'$ at every history in a $\mathcal{G}$-phase. By the weak Law of Large Numbers, this strategy meets the standards with the required probability, so long as Γ is sufficiently large. □

(Only if) Suppose not. Then there exists a sequence of σ_z which meet the standards defined by $\Gamma(z)$, $\chi(z)$ with probability approaching 1 as $z \to 0$, and for which $d_1(\mathbb{E}_\sigma[\ell(m,\omega)], q(m,\omega))$

[35] Such ε_R can always be found for $\varepsilon \leqslant 2V(\lambda')$.

stays bounded away from 0 by a constant, $c > 0$, for some (m, ω). Without loss, suppose $\mathbb{E}\left[\ell\left(m, \omega\right)\right] \leqslant q\left(m, \omega\right) - c$ (the other case is symmetric). By the definition of $\mathcal{S}_z$ we have

$$\Pr_{\sigma}\left[h'_{\Gamma(z)} \in \mathcal{S}_z\right] \leqslant \Pr_{\sigma}\left[\ell\left(m, \omega\right) \geqslant q\left(m, \omega\right) - \chi\left(z\right)\right].$$

But by Markov's inequality, the right-hand expression is bounded by

$$\Pr_{\sigma}\left[\ell\left(m, \omega\right) \geqslant q\left(m, \omega\right) - \chi\left(z\right)\right] \leqslant \frac{\mathbb{E}_{\sigma}\left[\ell\left(m, \omega\right)\right]}{q\left(m, \omega\right) - \chi\left(z\right)}.$$

As $z \to 0$ this bound is no more than $1 - \frac{c}{q(m,\omega)} < 1$ – a contradiction to σ_z passing with probability approaching 1 as $z \to \infty$. $\square$

Now we show that $d\left(\mathbb{E}_{\sigma}\left[\ell\right], q\right)$ can be bounded close to 0 only if the sender's strategy obeys (16), where $\varepsilon_S \to 0$ as $z \to 0$. Let $\ell\left(m, \theta, \omega\right)$ be the joint frequency of triple (m, θ, ω) in a $\mathcal{G}$-phase. Note first that if we can show that for any strategy σ and any (m, θ, ω)

$$\mathbb{E}_{\sigma}\left[\ell\left(m, \theta, \omega\right)\right] = \Pr\left[\omega \mid \theta, a\left(m\right)\right]\mu_0(\theta)\bar{\sigma}\left(m, \theta\right) \tag{18}$$

then Assumption 3 ensures us that $\mathbb{E}_{\sigma}\left[\ell\left(m, \omega\right)\right]$ and the time-average expected strategy are related by a *one-to-one* mapping of the form:

$$\mathbb{E}_{\sigma}\left[\ell\left(m, \omega\right)\right] = \sum_{\Theta} \Pr\left[\omega \mid \theta, a\left(m\right)\right]\mu_0(\theta) \cdot \bar{\sigma}\left(m, \theta\right).$$

The claim will then follow simply. We argue that (18) always holds by induction on Γ. To make this clear, we denote explicitly the dependence of the relevant likelihood on Γ by ℓ_{Γ}. Clearly, (18) holds for $\Gamma = 1$. To make the inductive step, assume (18) holds for all $\Gamma \leqslant T$. We show it also holds for $\Gamma = T + 1$. For any likelihood function, we can write:

$$\ell_{T+1}\left(m, \theta, \omega\right) = \frac{T}{T+1}\ell_T + \frac{1}{T}\ell_1. \tag{19}$$

Applying (19) to $\ell\left(m, \theta, \omega\right)$ and using the law of iterated expectations, we can write:

$$\mathbb{E}_{\sigma}\left[\ell_{T+1}\left(m, \theta, \omega\right)\right] = \frac{T}{T+1}\mathbb{E}_{\sigma}\left[\ell_T\left(m, \theta, \omega\right)\right] + \frac{1}{T+1}\mathbb{E}_{\sigma}\left[\mathbb{E}_{\sigma}\left[\ell_1\left(m, \theta, \omega\right) \mid h'_T\right]\right].$$

By the inductive hypothesis, both terms on the right-hand side can be rewritten using (18)[36]. But we can similarly interpret $\bar{\sigma}\left(m, \theta\right)$ as a likelihood function itself. Applying (19) in reverse yields the claim.

Since by Assumption 3, (18) implies that $\mathbb{E}\left[\ell\left(m, \omega\right)\right]$ and $\mathbb{E}\left[\bar{\sigma}\right]$ are related by a continuous bijection, for any $\eta_z > 0$ such that $d_1\left(\mathbb{E}_{\sigma}\left[\ell\right], q\right) \leqslant \eta_z$, there must exist a corresponding $\tilde{\varepsilon}_S > 0$ such that $d_1\left(\bar{\sigma}\left(m, \theta\right), \lambda'_{\mu'}\right) \leqslant \tilde{\varepsilon}_S$ for all $m \in \operatorname{supp}\{\lambda'\}$. Moreover, $\tilde{\varepsilon}_S \to 0$ as $\eta_z \to 0$ (alternatively, as $z \to 0$). Hence, we can indeed bound the sender's time average strategy appropriately, as $z \to 0$. $\square$

[36]Note that the history h'_T does not enter the relationship (18) for the period $T+1$

Hence for $(\varepsilon_S, \varepsilon_R)$ sufficiently small, we have established that SBS parameters exist such that best responses are mutually bounded in the compact convex strategy sets defined by (16). Thus, an equilibrium exists in this range, for each such $(\varepsilon_S, \varepsilon_R)$ (Debreu (1952), Glicksberg (1952), Fan (1952)). Taking limits as $\varepsilon_S, \varepsilon_R \to 0$, part 1 of the Theorem is immediate. Part 2 follows similarly from being able to trap the sender's time-average information structure arbitrarily close to $\lambda^\star$, and taking $\delta \to 1$. □

Online Appendix A

Proof of Lemma 1

Since there is a single long-run player, and (θ_t, ω_t) *(i)* enters stage payoffs at time t only, and *(ii)* is i.i.d. over time, Theorem 5.2 (Fudenberg and Levine (1994)) can be applied to conclude that it is without loss to restrict attention to Perfect Public Equilibria of the repeated cheap talk game.[37]

Consider some PPE strategy profile, where $\sigma_t : \underline{\mathcal{H}}_t \times \Theta \to M_t$ denotes the (mixed) behavior strategy of S at time t, and $a(\mu)$ denotes each receiver's optimal action, given posterior beliefs μ. Let receiver t's equilibrium beliefs given message m and public history $\underline{h}_t$ be $\mu_t(m, \underline{h}_t)$ and let $\lambda_t(\underline{h}_t)$ be the corresponding information structure provided to R_t. We construct a Direct Equilibrium that generates the same distribution over $(a_t, \theta_t, \omega_t)$ at each t as does the original PPE.

The construction is simple: at each public history, relabel any on-path message m with its induced equilibrium belief, $\mu = \mu_t(m, \underline{h}_t)$. If two distinct messages m_1, m_2 lead to the same posterior belief at history $\underline{h}_t$, but induce different continuation strategies, then at the end of round t we let a public randomization device (with appropriately chosen probabilities) determine which continuation is played. Finally, if a sender chooses an off-path message μ at history $\underline{h}_t$ in Direct Equilibrium, let R_t's corresponding (off-path) posterior belief be generated by a random draw corresponding to $\lambda_t(\underline{h}_t)$. Of course, this new strategy profile for the sender, and the induced beliefs, replicates the original outcomes, and satisfies all the criteria of a Direct Equilibrium.

We now show that it is without loss to restrict attention to message spaces such that $\mid M_t \mid \leqslant \mid \Theta \mid$ for all t. Suppose for a contradiction that for some equilibrium payoff $\mathbb{E}\left[V\left(\underline{h}_\tau, \theta_\tau\right)\right]$ of the sender and some history $\underline{h}_\tau$, the minimum number of messages in the sender's strategy compatible with obtaining $\mathbb{E}\left[V\left(\underline{h}_\tau, \theta_\tau\right)\right]$ in equilibrium is $|M'| = N' > \mid \Theta \mid$, where $M' = \text{supp } \lambda_t(\underline{h}_t)$. This strategy induces a N'-point distribution $\nu \in \Delta\left(\Delta\Theta\right)$ of posterior beliefs $\{\mu_\tau\left(m\right)\}_{m \in M'}$ over θ_τ and a corresponding distribution over receiver R_τ's actions, $a_\tau\left(\mu_\tau\left(m\right)\right)$,

[37] While our stage game is inherently multistage, this only serves to rule out strategies that are not best responses for short-run receivers at each history of the game. Since receiver's must best respond at each history in both PBE and Direct Equilibrium, essentially identical arguments to those in Fudenberg and Levine (1994) still go through.

where

$$a_\tau\left(\mu_\tau\left(m\right)\right) \in \arg\max_{a\in A} \mathbb{E}\left[u_R\left(a,\theta,\omega\right) \mid \mu_t\right] = \sum_{i=1}^{N} \mu_\tau^i \cdot \mathbb{E}\left[u_R\left(a,\theta^i,\omega\right) \mid \theta^i\right]$$

For this to be an equilibrium, it must be that for all $m_\tau \in \text{supp}\left(\sigma_\tau\left(\underline{h}_\tau,\theta\right)\right)$ and any $\tilde{m} \in M$,

$$\begin{aligned} V\left(\underline{h}_\tau,\theta_\tau\right): &= v\left(\mu_\tau\left(m_\tau\right),\theta_\tau\right) + \delta\mathbb{E}\left[V\left(\underline{h}_{\tau+1},\theta_{\tau+1}\right)\right] \\ &\geqslant v\left(\mu_\tau\left(\tilde{m}\right),\theta_\tau\right) + \delta\mathbb{E}\left[V(\left(\tilde{\underline{h}}_{\tau+1},\theta_{\tau+1}\right)\right] \end{aligned}$$

where $\underline{h}_{\tau+1} = \left(\underline{h}_\tau, m_\tau, a_\tau, \theta_\tau\right)$ and $\tilde{\underline{h}}_{\tau+1} = \left(\underline{h}_\tau, \tilde{m}, \tilde{a}_\tau, \theta_\tau\right)$. In particular, given any state θ_τ and messages $m_\tau, \tilde{m}_\tau \in \text{supp}\left(\sigma_\tau\left(\underline{h}_\tau,\theta\right)\right)$, we must have

$$v\left(\mu_\tau\left(m_\tau\right),\theta_\tau\right) + \delta\mathbb{E}\left[V\left(\underline{h}_{\tau+1},\theta_{\tau+1}\right)\right] = v\left(\mu_\tau\left(\tilde{m}_\tau\right),\theta_\tau\right) + \delta\mathbb{E}\left[V\left(\tilde{\underline{h}}_{\tau+1},\theta_{\tau+1}\right)\right]$$

Given any history, we define an equilibrium message $m^\theta \in M'$ to be *uniquely prescribed* at state θ if $supp\left(\sigma_\tau\left(h_\tau,\theta\right)\right) = \left\{m^\theta\right\}$. The set of all messages that are uniquely prescribed at some state $\theta \in \Theta_\tau$ is denoted M^Θ. We divide the set of equilibrium messages sent at history h_τ into two mutually exclusive and exhaustive sub-groups: those that are uniquely prescribed, $m \in M^\Theta$, and those that are not, $m \in M'/M^\Theta$.

Since $N' >\mid \Theta \mid$, there exists an $\tilde{m} \in M'$ and corresponding $\mu_\tau\left(\tilde{m}\right) \in \left\{\mu_\tau\left(m\right)\right\}_{m\in M'}$ that can be removed from the support such that remaining posteriors still satisfy Bayes' plausibility

$$\sum_{m_\tau\in M'/\{\tilde{m}\}} \alpha_{m_\tau}\mu_\tau\left(m_\tau\right) = \mu_0 \tag{20}$$

for some weights α_{m_τ} such that $\alpha_{m_\tau} \geqslant 0$, $\sum \alpha_{m_\tau} = 1$ (follows from Carathéodory's Theorem applied to the convex set, $\Delta\left(\Delta\Theta\right)$). By Proposition 1 in Kamenica and Gentzkow (2011), posteriors $\mu_\tau\left(\tilde{m}\right) \in \left\{\mu_\tau\left(m\right)\right\}_{m\in M'/\{\tilde{m}\}}$ can be sustained by a feasible signal structure with $N'-1$ distinct messages. Moreover, the message $\tilde{m}$ cannot be uniquely prescribed in any state $\theta \in \Theta$. Otherwise, there would exist some θ^i for which $\mu_\tau^i\left(m\right) = 0$, $\forall m \in M'/\left\{\tilde{m}\right\}$, while $\mu_0^i > 0$, violating (20). Therefore, $\tilde{m} \in M'/M^\Theta$ and for every state θ in which σ proscribes $\Pr\left(m_\tau = \tilde{m} \mid \underline{h}_\tau,\theta\right) > 0$, there exists another message m'_θ sent with positive probability in state θ.

Construct a new strategy $\sigma^\star$ which induces the distribution $\left(\alpha_{m_\tau}\right)_{m_\tau\in M'/\{\tilde{m}\}}$ over the posteriors $\left\{\mu_\tau\left(m\right)\right\}_{m\in M'}$ at history $\underline{h}_\tau$, and plays according to σ otherwise (this is feasible, by Proposition 1 of Kamenica and Gentzkow (2011)). For any $m \in M'/\left\{\tilde{m}\right\}$, the strategy continues to induce belief $\mu_\tau\left(m\right)$ at history h_τ and leaves continuation payoffs unchanged at $V\left(\underline{h}_\tau,\theta_\tau\right)$ thereafter (for any $\theta_\tau \in \Theta$). Moreover, this continuation payoff is well defined for each m since $\tilde{m}$ was never uniquely prescribed.

Therefore, strategy $\sigma^\star$ achieves the same payoffs for the sender from history $\underline{h}_\tau$ (due to indifference across all messages at that history), leaves payoffs otherwise unchanged at other histories, and involves only $N'-1$ messages sent at history $\underline{h}_\tau$. Therefore, it also does not affect incentive compatibility of equilibrium play at any prior history, $\underline{h}_t$, for $t < \tau$. It trivially does not affect the incentive compatibility of any history $\underline{h}_t$, for $t > \tau$. But this is a contradiction to N' as the minimum number of messages in any strategy consistent with $\mathbb{E}\left[V\left(\underline{h}_\tau, \theta_\tau\right)\right]$. Finally, we note that an essentially identical argument (with the appropriate adjustment of Carathéodory's Theorem) to the above establishes it is without loss to restrict attention to at most $N+1$-point distributions at each $\underline{h}^t$, for characterizing the set of attainable equilibrium payoff profiles for the sender and *all* receivers. $\square$

Online Appendix B: Support Material for Section 4

Here we develop the details behind the claims made in section 4.

Section 4.1

To be concrete, we develop arguments in the context of Example 1 with $l = -1$, $\mu_0 = \frac{1}{3}$, $p = 1$, and $\delta \to 1$. The threshold belief is $\underline{\mu} = 0.5$ and is induced by lying half the time when quality is low: $\lambda(0.5|\theta = l) = 0.5$. Under the assumption of random arrival, an average payoff of $\hat{v}(\mu_0) = \frac{2}{3}$ can be attained easily in equilibrium using a simple badge system with $\Gamma = 2$, $\beta = \infty$, and standards requiring no negative feedback in the second period of an evaluation. In this equilibrium the seller always lies in the first period of an evaluation and is honest in the second. Hence, it could not be an equilibrium if customers knew t — the first customer would never buy knowing the seller will lie whenever $\theta = l$. Nonetheless, when t is observable the platform could replicate the effect of random arrivals with a system that randomizes evaluation dates. We construct such a system below:

Proposition 6 *Consider Example 1 with $\mu_0 < \underline{\mu}$. Expected payoff $\hat{v}(\mu_0)$ is attainable using a badge system with stochastic evaluation dates when t is public.*

Proof. Consider the following badge system with random evaluation dates. At $t = 1$, the seller is given a rating $r_1 = \mathcal{G}$. At the outset, the system determines with equal probability whether the seller's $\mathcal{G}$ rating will be evaluated on odd, or on even, days – and this outcome is told *privately* to the seller. If an evaluation occurs in period t, the seller retains a $\mathcal{G}$ rating until the next evaluation if he avoids negative feedback from customer t. Otherwise, his rating is switched to $\mathcal{B}$ forever thereafter. At any t, the incoming customer observes t and the current value of r_t only.

We argue that the following strategy profile is an equilibrium for δ large enough. Moreover the seller's discounted average payoff converges to $\frac{2}{3}$ as $\delta \to 1$. In any evaluation, the seller always recommends 'buy' in the first period, irrespective of current θ_t, and recommends 'buy' in the second period if and only if the current $\theta_t = h$. Otherwise (if $r_t = \mathcal{B}$), the seller always recommends 'buy' again. Each receiver adopts the following strategy: if $r_t = \mathcal{G}$, the receiver obeys the seller's recommendations, while if $r_t = \mathcal{B}$ she does not buy. Checking equilibrium in this case is trivial. The only deviation that really needs checking is at those histories beginning in the second period of an evaluation. By the one shot deviation property, we need only check that the seller does not wish to deviate from honesty in the second period of an evaluation. This is the case if:

$$\mu_0 + \delta V^\star \geqslant 1 + \delta \mu_0 V^\star$$

where $V^\star = \frac{1+\delta\mu_0}{1-\delta^2}$ is his continuation payoff in the event of retaining a $\mathcal{G}$ rating. On rearrangement, this becomes

$$\delta \frac{1+\delta\mu_0}{1-\delta^2} \geqslant 1.$$

As $\delta \to 1$ the RHS of this expression becomes infinite. Hence, for δ large enough, we have verified the seller is best responding. Finally, applying l'Hôpital's rule verifies that the discounted average payoff $(1-\delta)V^\star \to \frac{2}{3}$. □

What if customers must be informed of the dates on which evaluations take place? Even here it turns out the platform can still replicate the required uncertainty, so long as it adopts more complex standards that depend on *full sequences* of outcomes rather than simple averages:

Proposition 7 *Consider Example 1 with $\mu_0 < \underline{\mu}$. Expected payoff $\hat{v}(\mu_0)$ is attainable using a badge system with complex standards and public deterministic evaluation dates.*

Proof. Consider the following standard. Let evaluation phases consist of T periods, where T is even. At the end of any evaluation, the seller is allowed to avoid a *permanent* suspension if his outcomes show either: *(i)* he lied only on even days, or *(ii)* he lied only on odd days. If he chooses the latter, he goes through a brief suspension. To be concrete, we should specify the length of the temporary badge suspensions imposed on a seller who adopts a strategy of lying on odd days. However, to avoid cumbersome notation we instead implement the punishment with a fixed probability q of losing the badge at the end of the evaluation. In particular, we choose

$$q = \frac{1-\mu_0}{\delta+\mu_0} \frac{\left(1-\delta^T\right)(1-\delta)}{\delta^T}.$$

Notice that q is decreasing in δ, with $q \to 0$ as $\delta \to 1$. Hence, there exists a threshold δ' such that q is indeed a well-defined probability, so long as $\delta \geqslant \delta'$.

We first argue that the strategy profile described in section 4 is an equilibrium, so long as $\delta \geqslant \max\{\delta', \delta''\}$, where $\delta'' < 1$ is the minimal value of x such that[38]

$$1 - x^T \leqslant x^T \cdot \frac{\mu_0 + x}{1+x}. \tag{21}$$

Verifying that customers best respond is trivial and hence omitted. Focus instead on the seller and let his continuation value from adopting the strategy proposed in the main text be $V^\star$. If he lies only on even days during an evaluation, he gets

$$\mu_0 + \delta + \delta^2\mu_0 + \cdots + \delta^{T-2}\mu_0 + \delta^{T-1} + \delta^T V^\star.$$

[38]This minimum is well-defined: the left (right) side of the inequality is continuous and decreasing (increasing) in δ. Finally, the inequality is violated at $x = 0$ and holds strictly at $x = 1$.

Similarly, if he lies on odd days he gets

$$1 + \delta\mu_0 + \delta^2 + \cdots + \delta^{T-2} + \delta^{T-1}\mu_0 + \delta^T(1-q)V^\star.$$

In order to be an equilibrium, the seller must be indifferent between these strategies. Using the maximum principle, it is easy to verify that $V^\star = \frac{\mu_0+\delta}{(1+\delta)(1-\delta)}$ and that indifference indeed holds for our choice of q. Moreover, any other deviation cannot be profitable for $\delta \geqslant \delta''$. In any evaluation, the seller's optimal deviation from his equilibrium strategy is trivially to lie in all periods. This is not profitable if

$$\frac{1-\delta^T}{1-\delta} \leqslant \delta^T V^\star$$

which on rearrangement is just (21), at $x = \delta$. Since $\delta \geqslant \delta''$ this is indeed satisfied. Finally, direct calculation for $\mu_0 = \frac{1}{3}$ verifies that $\lim_{\delta\to 1}(1-\delta)V^\star = \frac{2}{3}$. □

Section 4.2

Proposition 8 *Consider Example 1 with imperfect monitoring, $p < 1$. Using a blind sender review system, truth telling payoffs (μ_0, μ_0) are attainable in the limit as $\delta \to 1$. However, in any Perfect Public Equilibrium of such a system the sender cannot earn payoffs greater than μ_0.*

Proof. For the first claim, consider the following *blind sender* review systems. At each t, the incoming customer observes the complete history of feedback $\underline{h}_t$. Badges are again awarded to the seller based on evaluations of standards of length T, but now by the customers rather than a third party.[39] The important novelty here is that the seller no longer observes the individual feedback of past customers, but only the badges. Hence his history at $t > 1$ is $(r_\tau, \theta_\tau, m_\tau, a_\tau)_{\tau=1}^{t-1}$ and at $t = 1$ it is $r_1 = \mathcal{G}$. But this maps into the framework of Abreu et al. (1991). Hence we can apply their Proposition 6 to conclude that the (pure) strategy of truth telling can be supported with an expected punishment converging to 0 as the sender gets patient (and T gets large). This can be done by adopting standards in which the seller retains a $\mathcal{G}$ rating if and only if *every* product sold receives bad feedback.

For the final part of the Proposition, we focus on Perfect Public Equilibrium of *any* blind sender system (i.e. across all possible standards). As argued in the main text this is the relevant concept for understanding the 'reusable punishment' insight. But the logic of Fudenberg et al. (1990) can again be applied to this problem to show that the sender's payoff in any PPE is bounded by Proposition 1. In Example 1, this bound corresponds to the truth telling payoff. □

[39]For simplicity we assume here that if a sender ever receives a $\mathcal{B}$ rating, this is permanent. We allow for a PRD, so punishments can be made a probabilistic function of outcomes if necessary

Section 4.4

We briefly illustrate the wider applicability of the results of section 3.2 with an example:

Example 2 *Replace the stage game of Example 1 with the following moral hazard game. The customer chooses whether to buy,* $a \in \{B, N\}$*. The seller has no private information but chooses an effort level,* $e \in \{h, l\}$*. If the product is bought, the customer sees a noisy signal of effort* $\omega_t \in \{h, l\}$*, where* $p(\omega_t = e|e, a = B) > 0.5$*. Otherwise* $\omega_t = \varnothing$*. Payoffs are as follows:*

		Customer	
		B	N
Seller	h	$1, 1$	$0, 0$
	l	$3, -1$	$0, 0$

The sets of individually rational feasible stage payoffs and payoffs attainable in equilibrium are the same as in Example 1. From Theorem 2 we obtain the following.

Corollary 5 *Consider Example 2. Any individually rational, feasible payoff profile is attainable as an equilibrium with some SBS as* $\delta \to 1$*.*

The only important difference between Example 2 and 1 is what a customer learns about $\underline{h}_t$ and t before she acts. In 2 the seller's effort teaches her nothing, whereas in 1 his message provides some information. For instance, if a receiver observes a "used" message then she can infer she is not at a history at which the sender strictly gains from mis-selling. Since Example 2 does not suffer this inference problem, the argument underpinning Theorem 2 applies *a fortiori* to classic moral hazard problems too.Indeed, it would be a relatively simple extension of Theorem 2 to show this holds in a wide class of games with a long-run player facing a sequence of short-run players.