REPUTATION FOR A DEGREE OF HONESTY

Laurent Mathevet — *European University Institute*

David Pearce — *New York University*

Ennio Stacchetti — *New York University*

November 28, 2022

Abstract. Can reputation replace legal commitment for an institution making periodic public announcements? Near the limiting case of ideal patience, results of Fudenberg and Levine (1992) imply a positive answer in *value* terms, in the presence of a rich set of behavioral types. Little is known about equilibrium *behavior* in such reputational equilibria. Computational and analytic approaches are combined here to provide a detailed look at how reputations are managed. Behavior depends upon which of three reputational regions pertains after a history of play. These characterizations hold even far from the patient limit. Near the limit, a novel method of calculating present discounted values, stationary promising-keeping, helps establish a close connection between the reliability of the institution's reports and the Kamenica and Gentzkow (2011) commitment benchmark. It is striking that this connection still holds when the benchmark type is not available (in the set of behavioral types) to be imitated.

[†]The content of our earlier work, entitled "Reputation and Information Design," has been incorporated into this paper. Special thanks are due to Vishal Ashvinkumar and Rumen Kostadinov for exceptional research assistance. We are also grateful to Dilip Abreu, Ben Brooks, Drew Fudenberg, Ed Green, Emir Kamenica, Elliot Lipnowski, Daniel Monte, Stephen Morris and Debraj Ray for their helpful comments and conversations, as well as to seminar audiences at the University of California San Diego, the University of Edinburgh, the University of Toronto, Cornell University, Pennsylvania State University, the University of Illinois, the University of North Carolina, the University of Michigan, Rutgers University, Washington University in Saint-Louis, Columbia University, London School of Economics, New York University, the 2017 Latin American Meeting of the Econometric Society, the 2017 Becker Friedman Institute Economic Theory Conference and the 2020 North American Meeting of the Econometric Society.

Introduction

Repeated and dynamic games have become standard frameworks in which to analyze interactions that are not entirely governed by legal commitments. One of the most successful devices for narrowing the vast multiplicity of equilibria in standard repeated games involves introducing at least a small probability that the players may be types who act in certain predictable ways. Rational players can build reputations by imitating one of these types and make a lasting impact on expected future behavior. First studied by Kreps and Wilson (1982) and Milgrom and Roberts (1982), reputational models have been applied successfully to many economic problems.[1]

The exploding literature on Bayesian persuasion and information design[2] presents an ideal candidate for reputational analysis. One player, the sender, will learn some information of interest to another player, the receiver. The sender is assumed to be able to commit legally to the use of a particular protocol for sending messages to the receiver as a function of the state. While the most interesting cases involve the sender adopting a protocol that uses randomization, the range of circumstances in which it is plausible that one can commit legally to a random protocol is relatively narrow. But if a sender such as the IMF or the Federal Reserve Board periodically gets information of interest to the public, it might develop a reputation for filtering information in a certain way. Although it may not wish to be fully transparent, it could become known to use a certain degree of randomization to soften what it reveals.[3]

Analyzing the evolution of reputational dynamics for a random protocol is a hard problem. But for a very patient sender, there is a powerful result by Fudenberg and Levine (1992) (henceforth, FL) that, in this context, implies that the sender does virtually as well, in any perfect Bayesian equilibrium of the reputational dynamic game, as he would if he could commit legally to a random protocol.[4] The authors obtain this kind of result in impressive generality by studying the problem at a great distance: long run statistical arguments determine the asymptotic value, while nothing is established about how players behave in equilibrium.

In this paper, our goal is to understand the economics of reputational management and of asymptotic efficiency in dynamic information transmission. What happens in equilibrium? How does reputation affect the sender's reporting and the receiver's trust in the sender? Does the sender randomize between honest and dishonest reports?

[1]Backus and Driffill (1985), Fudenberg and Levine (1989), Bénabou and Laroque (1992), Abreu and Gul (2000), Mailath and Samuelson (2006), Abreu and Pearce (2007), Wolitzky (2012), Fanning (2016), and so on.

[2]Kamenica and Gentzkow (2011), Rayo and Segal (2010), Benoît and Dubra (2011), Bergemann and Morris (2016, 2019), Taneva (2019), Kamenica (2019), Mathevet, Perego and Taneva (2020), etc.

[3]After assuming his post as chairman of the Federal Reserve Board, Alan Greenspan told a subcommittee of the U.S. Congress, December 1987: "Since I've become a central banker, I've learned to mumble with great incoherence. If I seem unduly clear to you, you must have misunderstood what I said."

[4]See Section 3.1 and the Online Appendix.

How is he incentivized to send useful messages, and how can this be done efficiently? Little seems to be known about these questions,[5] because of the complexity of this stochastic, infinite-horizon strategic problem. To make progress, we study a canonical class of dynamic Cheap Talk games with binary states of nature, messages, and actions. This class is very stylized, but lends itself to a range of interpretations: an infectious disease expert wanting to promote healthy practices; a monopolist periodically selling a product whose quality is unknown to consumers; a company wanting its employees to work regardless of the state (as in Ely (2017)). The celebrated Kamenica and Gentzkow (2011) prosecuting attorney model is included in this class. The expert advisor model of Bénabou and Laroque (1992) presents a fascinating contrast, featuring opposing interests between sender and receivers in *every* state. More will be said about this as the paper unfolds, in particular in Section 5.6 and in the literature review.

We approach the problem from a computational perspective by first developing an APS (Abreu, Pearce and Stacchetti (1990)) algorithm for reputational models. Computations provide the building blocks for the analytical treatment and also a visual of FL's value convergence. A first numerical insight is that the largest equilibrium value is monotone in the sender's reputation. That is, reputation creates value. It seems intuitive that reputation is a valuable resource, but it is an endogenous property of the model that is more elusive than one might think, for the monotonicity of the lowest equilibrium value[6] might undermine that of the largest, by depriving the latter of necessary punishments. In all our computations, which span a range of discount factors and behavioral types, the largest equilibrium value is monotone in the sender's reputation.

In using computational methods to support the theoretical analysis of a dynamical system, our approach is similar to Phelan and Stacchetti (2001), Gorno and Iachan (2020) and to recent quantitative works in dynamic public finance (see Golosov et al. (2007) for a survey), asset pricing (e.g., Borovicka and Stachurski (2020)), and others. The literature review in Section 8 provides a more detailed account.

We characterize the sender-preferred equilibrium in a tripartite division of the reputation space, even for discount factors away from one. Behavior is qualitatively uniform within regions, but incentives are provided in different ways across regions. When his reputation is very high, the sender exploits it by lying with certainty and the receiver nonetheless trusts any recommendation she receives. In this situation, which is labelled Region 3, marginal returns to investing further have fallen enough that it is time to "cash in".[7] Outside of Region 3, the sender is faced with the choice

[5] Cripps, Mailath and Samuelson (2004) teaches us that the sender will eventually lose his reputation to be trustworthy (in equilibrium) and, thus, that continuation play in the ultra-long run is a Nash equilibrium of the standard repeated game.

[6] The computational results indeed indicate that in most parameter ranges, the lowest equilibrium value increases with reputation.

[7] Anecdotal evidence suggests that some firms have taken liberties with their usual standards after reaching peaks of popularity, such as Toyota (safety recalls) and Volkswagen ("dieselgate").

of investing or disinvesting in his reputation, which he resolves stochastically. In particular, in Region 1, the sender's probability of investing in his reputation (by telling the truth) is decreasing in his reputation level and governed by a simple monotonic formula. Numerically, this monotonicity property holds beyond Region 1: with rare exceptions, the probability of reputational investment is lower at higher reputation values.

One of the most notable phenomena happens in Region 2, where the sender's incentives are met by using *optimal* continuation values. Whenever the sender randomizes, the continuation values need to be adjusted to maintain indifference, which can induce a per period cost that does not arise in the static problem with legal commitment. In the standard repeated game, this cost is often so severe that there is no benefit at all to randomizing compared to telling the truth (see Fudenberg, Kreps and Maskin (1990)). In the reputational model, however, the asymptotic value result of FL implies that this inefficiency is avoided almost entirely. That this cost can be circumvented is almost troubling, because it follows from a simple value recursion that would seem absolutely unavoidable. Region 2 behavior offers a resolution of this puzzle as indifference is achieved without destroying surplus, by exploiting the dynamic (as opposed to strictly repeated) nature of the game.

Under the monotonicity assumption that a sender with better reputation exploits it by lying with greater probability, a behavioral convergence result obtains from our equilibrium structure: for any behavioral type, positive ϵ and number N, there exists a discount factor above which, if the initial reputation (which is the receiver's initial belief that the sender is the behavioral type) lies in $[\epsilon, 1-\epsilon]$, then play by the sender and the receiver in the first N periods almost coincides with the stage-game equilibrium behavior. That is, it coincides with the static Bayesian persuasion behavior. As the behavioral type approaches the Bayesian persuasion ideal, this equivalence implies that (near-) Bayesian persuasion behavior carries so much weight in discounted terms that dynamic payoffs approach the static, concave envelope payoffs (Kamenica and Gentzkow (2011)). This is illuminating given the results of FL and Cripps, Mailath and Samuelson (2004): the former suggest an imitation strategy that cannot be an equilibrium and the latter predicts repeated Nash behavior in the ultra long-run which cannot determine discounted average payoffs (in light of the former).

The analysis in Sections 5 and 6 can be interpreted as "subgames" of a larger game, considered in Section 7, in which the sender first chooses which reputational posture to imitate. It might seem obvious that the sender should choose a behavioral type as close as possible to the static commitment ideal. But this is wrong. Interestingly, for any fixed discount factor, no matter how high, if the behavioral type is too close to that static ideal, the sender will do poorly except at extremely high reputation values: his payoff is bounded above by the best payoff in the game in which he has no behavioral type to imitate and hence no reputation to build. The intuition and formal details are included in Section 7, complemented by numerical examples.

An additional objective of this paper is to provide a dynamic foundation for the commitment assumption underlying Bayesian persuasion and information design. Although payoff convergence (Proposition 1) is presented in the context of our model, this result holds much more generally in virtue of FL and provides a justification for the standard Bayesian persuasion model. In the Online Appendix, we state the general claim, contrast it with the *failure* of payoff convergence in the standard repeated game without reputational concerns (see Fudenberg, Kreps and Maskin (1990) and, in the context of repeated cheap talk, Best and Quigley (2017)), and discuss the apparent paradox this raises.

The paper is organized as follows. Section 2 lays out the model. Section 3 presents the value convergence result. After Section 4 describes the computational method, Section 5 provides computational evidence to support monotonicity of the best equilibrium and describes the equilibrium structure theoretically. Section 6 presents the behavioral convergence result. The development of that result features a new technique that we call stationary promise-keeping. Section 7 touches on the model with many types. Section 8 provides a literature review and Section 9 concludes. Most of the proofs are in the Appendix. The computer codes that generate our numerical results and figures can be found at `https://tinyurl.com/yc6eantw`.

2. The Model

A sender communicates information to a receiver about an i.i.d state at discrete time periods $t \in \{0, \dots\}$. The sender can be either a rational type s_R or a behavioral type s_B (as described below), which is private information to the sender. The sender is behavioral with prior probability β_0 and rational with the residual probability.[8]

The receiver begins period t with a belief $\beta_t \in [0,1]$ that the sender's type is s_B. This belief represents the sender's reputation. State θ_t is drawn from $\Theta = \{\ell, h\}$ according to the distribution $\mu_0 \in \Delta\Theta$. When there is no confusion, we simply write $\mu_0 = \mu_0(h)$. The sender observes the state realization and then sends a message $m_t \in M = \{L, H\}$ to the receiver. Denote the rational sender's strategy at t by $\pi_t : \Theta \to \Delta(M)$ where $\pi_t(\cdot|\theta)$ is the distribution of messages given $\theta_t = \theta$ (π_t may also depend on the observed history of play before t). Denote the behavioral type's strategy by $\pi_B : \Theta \to \Delta(M)$ and assume $\pi_B(H|h) = 1$. Since s_B is always truthful in the high state by assumption, we use π_B to denote $\pi_B(H|\ell)$ when there is no confusion.

Later in period t, upon receiving message m_t, the receiver updates her belief and chooses an action. Let $\mu_t(\cdot|m_t)$ be the receiver's posterior distribution about θ_t (which depends on β_t, because the sender's identity matters to interpret his message, and on

[8]Before the game starts, Nature draws a number β_0 uniformly from $[0,1]$ and then draws the sender's type such that he is rational with probability β_0. Having Nature draw the prior at the beginning of the game is just a convenient device to study all the games beginning at different priors β_0 at once. The uniform distribution plays no role at all in the analysis.

π_t (known in equilibrium) and π_B). Given her utility function

$$u(\theta, a) = \begin{cases} 1 & \text{if} \quad (\theta, a) \in \{(H, h), (L, \ell)\} \\ 0 & \text{otherwise,} \end{cases}$$

the receiver chooses action

$$a_t^*(m_t) \in \operatorname*{argmax}_{a \in A} \sum_{\theta \in \Theta} \mu_t(\theta | m_t) u(\theta, a).$$

Whereas the receiver always wants to match the current state with her action, the rational sender wants her to always choose the high action, which is captured by the following sender's payoff

$$v(\theta, a) = \mathbb{1}_{\{a=H\}}.$$

The sender maximizes her average discounted expected continuation value, where $\delta \in (0, 1)$ is the discount factor. Throughout the paper, we assume:

Assumption: $\mu_0 < 1/2$.

Assumption (Bounded discounting): $\delta \geq 1/(1 + \mu_0)$.

Note that the receiver always plays a myopic best response to her beliefs in each period. Anonymity in the large is a standard justification for such a behavior. Because any receiver among the public is anonymous from the view of the long-run institution, her deviations are not detectable and hence she plays myopically.

This model admits a range of interpretations as long as the receiver (a worker, buyer, citizen, etc.) finds it worthwhile to choose H (high effort, purchase, precautions) only in the high state, whereas the sender (an employer, seller, medical expert) would like her to choose H regardless of the state. For the sake of illustration, think of a health agency that periodically monitors the state of an infectious disease and that can issue a warning to the public in case the risk of infection is high. Since strict health practices are costly, the public prefers to adopt such behavior only if the risk is high.

Assumption (Trusted behavioral type): $\pi_B < \frac{\mu_0}{1-\mu_0}$.

This standard assumption ensures that if the receiver were certain that the sender is s_B, then given her payoffs she would want to follow his advice because $\mathbb{P}[\theta_t = h | H] > 1/2$.

At the end of period t, the state and the sender's message in that period, (θ_t, m_t), become commonly known (to all future players), which yields the next reputation

$$\beta_{t+1} = \beta_{m_t, \theta_t}(\beta_t, \pi_t) \equiv \frac{\beta_t \pi_B(m_t|\theta_t)}{\beta_t \pi_B(m_t|\theta_t) + (1 - \beta_t)\pi_t(m_t|\theta_t)}.$$

Let $V(\beta)$ be the set of PBE continuation values for the rational sender given the current receiver's belief β. The map $V : [0, 1] \rightrightarrows \mathbb{R}$ is the equilibrium value correspondence, which also depends on the discount rate δ and sometimes we write $V(\beta, \delta)$

to make this explicit. A public randomization device ensures that $V(\beta)$ is a convex set: at the beginning of every period, before the state is realized, the players publicly observe the outcome of a draw from a uniform distribution in $[0,1]$.

3. Asymptotic Efficiency

This Section translates the FL asymptotic value result for a long-lived player, to the particular dynamic sender-receiver game of interest here.

Assuming that the sender is known to follow strategy π_B, let $\mu^{\pi_B}(\cdot|m)$ be the receiver's Bayesian posterior belief about the state given m[9] and

$$\begin{aligned} v^{\pi_B} = \min \quad & \sum_{\theta\in\Theta}\sum_{m\in M} \mu_0(\theta)\pi_B(m|\theta)v(\theta,a_m) \\ \text{s.t.} \quad & a_m \in \operatorname*{argmax}_{a\in A} \sum_{\theta\in\Theta} \mu^{\pi_B}(\theta|m)u(\theta,a) \end{aligned}$$

be the sender's payoff when the receiver chooses the sender-worst best reply to π_B. Assuming $\mu_0(h) < 1/2$, $v^* = 2\mu_0$ is the optimal value that a sender with commitment can attain in the Bayesian persuasion problem[10] by following $\pi(H|h) = 1$ and

$$\pi(H|\ell) = \pi^* := \frac{\mu_0}{1-\mu_0},$$

as computed by standard techniques (e.g., Kamenica and Gentzkow (2011)).

Proposition 1. (Value Convergence) *For each $\beta \in (0,1)$ and $\epsilon > 0$, there exists $\underline{\delta} < 1$ such that $V(\beta,\delta) \subseteq [v^{\pi_B} - \epsilon, v^* + \epsilon]$ for all $\delta \geq \underline{\delta}$.*

In all PBE, the rational sender receives at least v^{π_B} and at most his Bayesian persuasion value, both approximately. This follows from FL. As a result, if π_B is close to π^*, then a patient sender is virtually guaranteed in all PBE the same value as he would have under full commitment power, because $\lim_{\pi_B\uparrow\pi^*} v^{\pi_B} = v^*$.

4. Computational Method

This Section describes how to compute the equilibrium value correspondence for the sender, for any values of the parameters (discount factor δ and degree of dishonesty π_B of the behavioral type), by developing analogues of the APS (1990) results and algorithm for reputational models. Readers more interested in the numerical and analytical results can comfortably skip to Sections 5 and 6.

---

[9] For all θ and m, $\mu^{\pi_B}(\theta|m) = \frac{\pi_B(m|\theta)\mu_0(\theta)}{\sum_{\theta\in\Theta}\pi_B(m|\theta)\mu_0(\theta)}$.

[10] The case $\mu_0 \geq 1/2$ is trivial, because disclosing no information leads the receiver to choose action H and hence $v^* = 1$.

We describe the method in its general form and all the results in this Section (4.1 and 4.2) hold for any finite sets Θ, A and M, any finite set of sender's types $\hat{\Pi} = \{s_R, s_B^1, \ldots, s_B^K\}$ and arbitrary payoffs for the sender and the receiver. Each type s_B^n is characterized by an informational map $\pi_B^n : \Theta \to \Delta(M)$.

For a fixed δ, the PBE correspondence V may be computed iteratively, starting from the set of feasible payoffs of the sender. Let $W^0(\beta) = [\underline{v}, \overline{v}]$ for all β, where $\underline{v} = \min v$ and $\overline{v} = \max v$. Given W^k, the iterative step constructs $W^{k+1}(\beta)$ as the set of values of all admissible pairs at β, with continuation payoffs respecting W^k. Below we first introduce admissibility and then show that W^{k+1} inherits from W^k the following properties: W^{k+1} has a closed graph, and for each β, $W^{k+1}(\beta)$ is a nonempty and convex set. Since $\{W^k\}$ is a decreasing sequence of nonempty compact sets, $W^\infty(\beta) = \bigcap_{k\in N} W^k(\beta)$ is nonempty, by the finite intersection property. A self generation argument shows that $W^\infty = V$.

4.1 Self Generation

Assume that in the current period the receiver has belief $\beta \in \Delta(\hat{\Pi})$ and expects the sender to play strategy π. After receiving message m, the receiver updates her beliefs about the state and chooses a myopic (possibly mixed) best reply in $\mathrm{BR}_m^R(\beta, \pi)$. Denote $\mathrm{BR}^R(\beta, \pi) = \prod_m \mathrm{BR}_m^R(\beta, \pi)$.

If the sender expects the receiver to play strategy $\alpha \in \Delta(A)^M$ in the current period and continuation value $w_{m,\theta}$ after sending m in state θ, then his total expected value for strategy π is

$$E(\pi, \alpha, w) = \sum_{\theta\in\Theta} \mu_0(\theta) \sum_{m\in M} \pi(m|\theta) \left[(1-\delta) \sum_{a\in A} v(\theta, a)\alpha_m(a) + \delta w_{m,\theta}\right]$$

where $w = (w_{m,\theta})_{m,\theta}$. The sender's best-reply set is then

$$\mathrm{BR}^S(\alpha, w) = \mathrm{argmax}\,\{E(\pi, \alpha, w) | \pi : \Theta \to \Delta(M)\}.$$

Let $\mathcal{C}$ be the set of correspondences $W : \Delta(\hat{\Pi}) \rightrightarrows [\underline{v}, \overline{v}]$ having a compact graph and such that $W(\beta)$ is a nonempty convex set for each β. We will often use W to denote the graph of the correspondence, so that $W \subseteq W'$ means $W(\beta) \subseteq W'(\beta)$ for all β.

Definition. *Given $W \in \mathcal{C}$, the tuple (π, α, w) is admissible for W at $\beta \in \Delta(\hat{\Pi})$ if $w_{m,\theta} \in W(\beta_{m,\theta}(\beta, \pi))$ for each $(m, \theta) \in M \times \Theta$ and*

$$\pi \in BR^S(\alpha, w) \quad \textit{and} \quad \alpha \in BR^R(\beta, \pi).$$

A tuple (π, α, w) is admissible if (π, α) is a *static* Nash equilibrium of a game whose payoffs are given by our per-period payoffs augmented with continuation payoffs w.

Define correspondence $B(W)$ to be the set of sender's admissible payoffs given W,

$$B(W)(\beta) = \{E(\pi, \alpha, w) \mid (\pi, \alpha, w) \text{ is admissible for } W \text{ at } \beta\},$$

and define $\tilde{B}(W)$ as $\tilde{B}(W)(\beta) = \mathrm{co}(B(W(\beta)))$ for each β. We convexify the image of B and, indirectly, the PBE payoffs, in virtue of our public randomization.

Definition. *W is self-generating if $W \subseteq \tilde{B}(W)$.*

Proposition 2. *If $W \in \mathcal{C}$ is self-generating, then $\tilde{B}(W) \subseteq V$.*

If a payoff correspondence is self-generating, so that each value can be generated by an admissible tuple whose continuation payoff can also be generated by an admissible tuple and so on, then it must be a subset of the PBE correspondence.

4.2 Algorithm and Existence

The algorithm applies the $\tilde{B}$ map iteratively to a set that is originally large enough to include all possible PBE payoffs. In particular, starting from the set of feasible payoffs $W^0 = \Delta(\hat{\Pi}) \times [\underline{v}, \overline{v}]$, which is in $\mathcal{C}$, the $\tilde{B}$ map reduces its size without ruling out any PBE payoff:

$$V \subseteq \tilde{B}(W^0) \subseteq W^0.$$

Since the $\tilde{B}$ map is monotonic, iterative applications from W^0 keep on shrinking the set without ever excluding any PBE. Formally, let $W^{k+1} = \tilde{B}(W^k)$, $k \geq 0$. The sequence $\{W^k\}$ is decreasing: $W^{k+1} \subseteq W^k$ for all $k \in \mathbb{N}$. Let

$$W^\infty = \lim_{k\to\infty} W^k = \bigcap_{k \in N} W^k.$$

Importantly, $\tilde{B}(W)$ is nonempty valued when $W \in \mathcal{C}$, which is crucial for non-vacuous convergence of the iterative process.

Lemma 1. *If $W \in \mathcal{C}$ then $\tilde{B}(W) \in \mathcal{C}$.*

The crucial step in the proof of Lemma 1 is demonstrating, for each possible β, the existence of a tuple admissible with respect to W. While this is similar to establishing the existence of Nash equilibrium in a static game, there is the added complication of finding continuation payoffs from the appropriate value sets at respective updated reputations. The proof essentially adds a dummy player whose strategy selects the continuation payoffs. In the resulting three player game, we let Kakutani's fixed point theorem find suitable continuation payoffs along with strategies for the sender and the receiver.

Proposition 3. *$W^\infty \in \mathcal{C}$ and $W^\infty = V$*

As an immediate corollary, a PBE exists because $V \in \mathcal{C}$.

5. Equilibrium Behavior in Three Reputation Regions

This Section combines theoretical analysis and numerical computation to reveal the workings of reputation management for general discount factors. Reputation space can be divided into three regions with distinctly different characters. For reasons to become clear below, one can think of behavior in the three regions as inefficient provision of incentives for a degree of honesty, efficient provision of those incentives, and exploitation of reputation, respectively. The challenging exercise of characterizing limiting behavior as the Sender approaches ideal patience is postponed until Section 6.

Our trusted-type assumption motivates the sender to build his reputation for being the behavioral type. It is intuitive that such a reputation is valuable: a high reputation makes one more credible. But this is not straightforward theoretically. Suppose that the worst PBE value for the sender is increasing in reputation (this is usually the case – see 5.1 below). Then, some punishments become unavailable as reputation grows, so it may be harder to maintain good behavior by the sender, and hence trust by the receiver, at higher reputation levels. In that case, the best PBE value might not always be increasing with reputation.

Fortunately, that perverse case never arises in our grid search on parameter space. After a quick description of the grid of parameters used in the numerical work, we begin the analysis, introducing numerical results where they are needed.

5.1. Grid Search

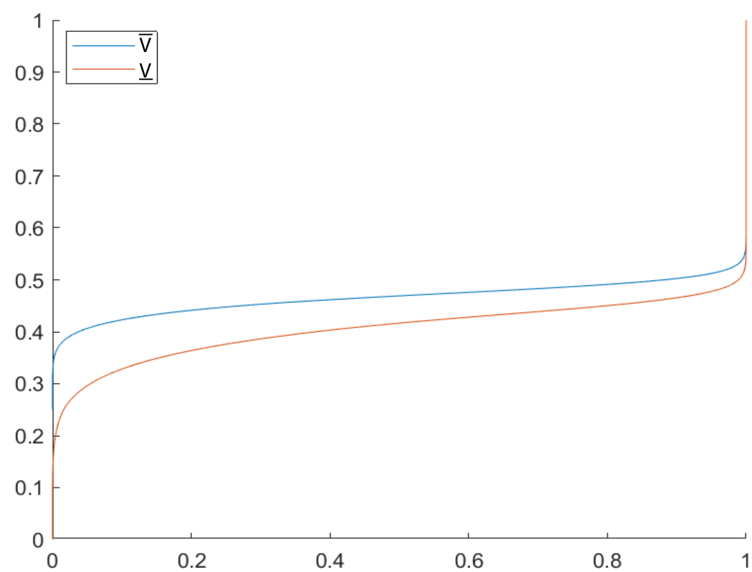

Figure 1: Value Correspondence for $\mu_0 = 0.25$, $\pi_B = 0.2$ and $\delta = 0.98$.

For all (μ_0, δ) in the grid $\{.1, .15, \dots, .45\} \times \{.9, .905, \dots, 0.995, 0.999\}$ and π_B in the grid $\{0.02, 0.04, \dots, \bar{\pi}_B(\mu_0)\}$, where $\bar{\pi}_B(\mu_0) < \mu_0/(1-\mu_0)$,[11] the modified APS

[11] More precisely, $\bar{\pi}_B(\mu_0) = 0.11, 0.17, 0.24, 0.33, 0.42, 0.53, 0.66, 0.81$ for $\mu_0 = 0.1, 0.15, \dots, 0.45$, respectively.

algorithm from Section 4 was used to compute the entire PBE value correspondence, mapping reputation levels β into the set of PBE values of the dynamic game having that initial reputation. Figure 1 graphs the results for $\mu_0 = 0.25$, $\pi_B = .2$ and $\delta = .98$.

One property was found to hold at all points in the grid search: defining $\overline{V}(\beta, \delta) = \max V(\beta, \delta)$,

Numerical Property 1. For each δ, $\overline{V}(\beta, \delta)$ is everywhere strictly increasing in β.

The results derived analytically for the remainder of the paper rely on Property 1. We will maintain this property (strict monotonicity of the upper boundary of the value correspondence) as an assumption throughout.

Next, define $\underline{V}(\beta, \delta) = \min V(\beta, \delta)$. The lower boundary of the value correspondence is not as universally monotone as the upper boundary. The numerical computations reveal that for some values of the discount factor, $\underline{V}(\beta, \delta)$ is not strictly increasing in β. But for relatively patient senders, the lower boundary is also increasing:

Numerical Property 2. For $\delta \geq 0.925$, $\underline{V}(\beta, \delta)$ is everywhere strictly increasing in β.

Once the structure of equilibrium behavior has been developed in a series of analytical results, further numerical patterns will be reported and interpreted in 5.3.

5.2 Equilibrium Value Correspondence

To get a sense of the general shape of V, as in Figure 1, consider what happens at the extreme reputations $\beta = 0$ and $\beta = 1$. Given our trusted-type assumption, the sender can get an infinite unbroken string of H when $\beta = 1$, because the receiver is certain that the sender is behavioral. Therefore, $V(1) = \{1\}$. Things are only slightly more complicated at $\beta = 0$. Here, the receiver is convinced that she faces a rational sender so that $V(0)$ is the equilibrium value set in the standard repeated game. Any "babbling equilibrium", where no information is conveyed and the receiver chooses L, gives the sender $\underline{V}(0) = 0$. Proposition A.2 in the Appendix shows that $\overline{V}(0) = \mu_0$, which corresponds to the truth-telling equilibrium: the sender reveals the state in each period, the receiver plays accordingly, and deviations are punished by babbling forever. Therefore, public randomizations yield $V(0) = [0, \mu_0]$.

5.3 Upper Boundary Behavior

We characterize equilibrium behavior with initial value on the upper boundary by showing that the reputation space can be organized into three regions.

For any $\beta \in [0, 1]$, in the first period of a PBE with value $\overline{V}(\beta)$, let $\pi(m|\theta, \beta)$ be the probability that the sender chooses message m in state θ and $\alpha(\beta)$ be the probabiity that the receiver accepts the recommendation H.

For each outcome (m, θ) of the first period, $w_{m,\theta} \in V(\beta_{m,\theta})$ is the associated continuation value, where $\beta_{m,\theta}$ is the reputation after (m, θ), derived in Section 2. Then

$$\overline{V}(\beta) = \mu_0 v_h + (1 - \mu_0) v_\ell \text{ where, for } \theta = \ell, h, \qquad \textbf{(Value Recursion)}$$
$$v_\theta = \pi(H|\theta, \beta)[\alpha(\beta)(1 - \delta) + \delta w_{H,\theta}] + \pi(L|\theta, \beta)\delta w_{L,\theta}$$

If $0 < \pi(H|\theta, \beta) < 1$, then it must be optimal for the sender to recommend H or L after observing θ, so he must be indifferent:

$$\delta w_{L,\theta} = (1 - \delta)\alpha(\beta) + \delta w_{H,\theta}.$$

Our first few lemmas describe basic properties of equilibrium behavior, from which the basic law of motion of reputation follows.

The rational sender's myopic incentives always tempt him to recommend H, and his only motivation for saying L is to improve his reputation. In the high state, however, saying L ruins the sender's reputation and so he always prefers saying H:

Lemma 2. (Honesty in high state) $\pi(H|h, \beta) = 1$ *for all* β.

By Lemma 2, there is no reputation updating in state h: $\beta_{H,h} = \beta$ (and $\beta_{L,h} = 0$ off path). Since $\pi(H|h, \beta) = 1$, we focus on $\pi(H|\ell, \beta)$ in what follows and denote it by $\pi(\beta)$.

When the state is low, both the rational and the behavioral type lie randomly, but the rational type is the more likely to say H:

Lemma 3. (Opportunistic dishonesty in low state) $\pi(\beta) > \pi_B$ *for all* $\beta > 0$.

Thus, considering a history after which the sender has reputation β and equilibrium strategy π, if the sender says H in the low state, his reputation falls to $\beta_{H,\ell}(\beta, \pi) < \beta$, whereas it rises to $\beta_{L,\ell}(\beta, \pi) > \beta$ if he says L.

Next, we show that on the upper boundary, the receiver always has at least a weak incentive to follow the sender's advice. By Lemma 2 and Bayes' rule, let

$$\mu(\pi, \beta) = \frac{\mu_0}{\mu_0 + (1 - \mu_0)(\beta\pi_B + (1 - \beta)\pi(\beta))}$$

be the receiver's posterior belief that $\theta = h$ given message H. Let

$$\bar{\pi}(\beta) := \frac{\pi^* - \beta\pi_B}{1 - \beta} \quad \text{for} \quad \beta \in [0, 1] \tag{1}$$

and note $\mu(\bar{\pi}(\beta), \beta) = 1/2$. That is, if in equilibrium the sender recommends H in state ℓ with probability $\bar{\pi}(\beta)$, the receiver is indifferent about following this recommendation.

Lemma 4. (Trustworthiness) *$\pi(\beta) \leq \bar{\pi}(\beta)$ for all β.*

Hence, following the sender's advice is always one of the receiver's best responses.

5.3.1. Region 3: Exploitation

When the sender's reputation is extremely high, he is no longer willing to invest further by admitting that $\theta = \ell$. He says H with probability 1. Nonetheless, because β is so high, the receiver always follows the sender's advice. Implicitly define $\bar{\beta}$ by

$$\overline{V}(\bar{\beta}) = \frac{\delta - \mu_0(2\delta - 1)}{1 - \delta\mu_0},$$

which is defined uniquely due to the strict monotonicity of $\overline{V}$. The next lemma shows that $\pi(\beta) = 1$ for all reputations in $[\bar{\beta}, 1]$, which we call Region 3.

Lemma 5. (Exploiting high reputations) *$\pi(\beta) = 1$ if and only if $\beta \geq \bar{\beta}$. Moreover, $\alpha(\beta) = 1$ for all $\beta \geq \bar{\beta}$.*

When the sender chooses $\pi(\beta) = 1$, he could obtain a reputation of 1 by deviating once and recommending L instead, allowing him to induce H forever after. This deviation would yield value $v_\ell = \delta$ in the low state. The smallest $\overline{V}(\beta)$ which satisfies the above value recursion while offering $v_\ell = \delta$, and which therefore supports $\pi(\beta) = 1$, is $[\delta - \mu_0(2\delta - 1)]/[1 - \delta\mu_0]$.

5.3.2. Region 2: Efficient Incentives

The interval $[0, \overline{\beta})$ is divided into Regions 1 and 2. In this interval, Lemmas 3 and 5 imply that the rational sender strictly randomizes in the low state. Since he must be indifferent, his payoff $\overline{V}(\beta)$ can be evaluated equivalently by what he gets by recommending L and moving his reputation to the right (right promise keeping):

$$\overline{V}(\beta) = \mu_0[(1-\delta)\alpha(\beta) + \delta\overline{V}(\beta)] + (1-\mu_0)\delta w_{L,\ell}$$

or lying and moving his reputation to the left (left promise keeping):

$$\overline{V}(\beta) = \mu_0[(1-\delta)\alpha(\beta) + \delta\overline{V}(\beta)] + (1-\mu_0)[(1-\delta)\alpha(\beta) + \delta w_{H,\ell}].$$

Equating these two and simplifying gives the incentive constraint

$$\delta[w_{L,\ell} - w_{H,\ell}] = (1-\delta)\alpha(\beta). \qquad \textbf{(IC)}$$

Lemma 6. *For all $\beta \in [0,1]$, the truthtelling continuation values are efficient, that is, $w_{H,h} = \overline{V}(\beta)$ and $w_{L,\ell} = \overline{V}(\beta_{L,\ell}(\beta, \pi(\beta)))$, and whenever the sender chooses $\pi(\beta) < \bar{\pi}(\beta)$, so is the remaining continuation value, that is, $w_{H,\ell} = \overline{V}(\beta_{H,\ell}(\beta, \pi(\beta)))$.*

As the sender increases $\pi(\beta)$ (the probability of false recommendation in the low state), $\beta_{L,\ell}$ increases and $\beta_{H,\ell}$ decreases (by Lemmas 2 and 3). It follows that $\overline{V}(\beta_{L,\ell}) - \overline{V}(\beta_{H,\ell})$ increases, so that $\pi(\beta)$ is increased until that difference equals $(1-\delta)/\delta$. If this can be done while $\pi(\beta) < \bar{\pi}(\beta)$, then not only does the receiver surely follow the sender's recommendation, $\alpha(\beta) = 1$, which is good for the sender, but (as proven in Lemma 6) all continuation values can be picked from the upper boundary.

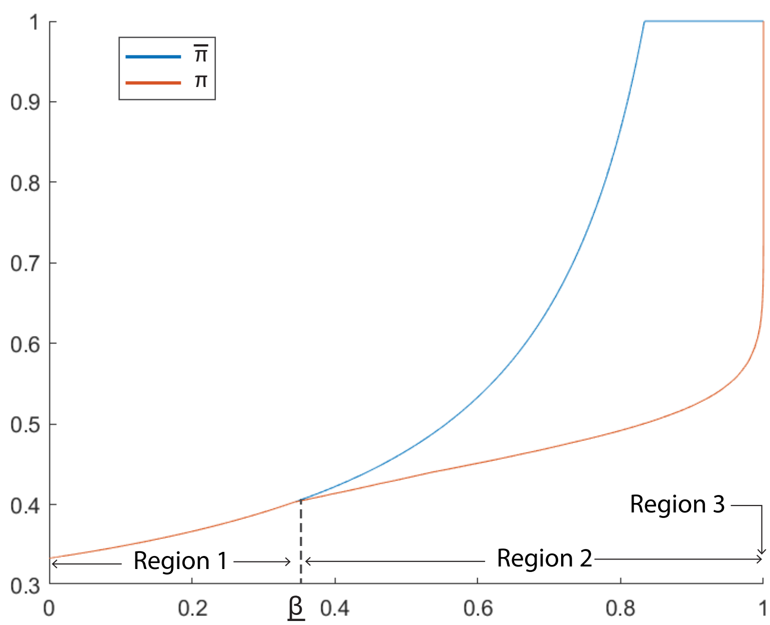

Figure 2: Sender's Equilibrium Strategy ($\mu_0 = 0.25$, $\pi_B = .2$ and $\delta = .98$)

Region 2 comprises those reputations β for which the sender can be incentivized in the most efficient way:

– the sender chooses $\pi(\beta) < \min\{\bar{\pi}(\beta), 1\}$ (hence the receiver *strictly* prefers to follow the sender's advice, $\alpha(\beta) = 1$). Let

$$\underline{\beta} = \inf\{\beta : \pi(\beta) < \min\{\bar{\pi}(\beta), 1\}\}$$

be the infimum of Region 2.

– all continuation values, especially $w_{H,\ell}$, are on the upper boundary.

Interestingly, since $\alpha(\beta) = 1$ in Region 2, α is *not* chosen to achieve the sender's indifference. Instead, the sender randomizes to keep *himself* indifferent.

5.3.3. Region 1: Inefficient Incentives

In Region 1, both the sender and the receiver are indifferent about their actions. To make the receiver indifferent, the sender uses $\pi(\beta) = \bar{\pi}(\beta)$. Two subregions describe alternative ways of satisfying the sender's incentives:

– **Region 1a**: the receiver always follows the sender's recommendations: $\alpha(\beta) = 1$, and $w_{H,\ell}$ is lowered from $\overline{V}(\beta_{H,\ell}(\beta, \pi(\beta)))$ so that (IC) holds (by increasing the sender's incentives to be honest).

– **Region 1b**: the receiver only follows recommendations stochastically: $\alpha(\beta) < 1$, and the continuation value lies on the lower boundary: $w_{H,\ell} = \underline{V}(\beta_{H,\ell}(\beta, \pi(\beta)))$.

In comparison to Region 2, inefficient payoffs (below the upper boundary) must be employed, both in 1a and 1b, as punishments for false reports of H in state ℓ. If such punishments are still not enough to provide the needed incentives for truth-telling, even more highly inefficient devices are needed (in 1b) by making the receiver sometimes disobey the sender's suggestions. Indeed, setting $\alpha(\beta) < 1$ discourages dishonesty by sometimes ignoring the "advice" H. This is a poorly targeted incentive, as it punishes message H even when that message is truthful! (By comparison, dropping $w_{H,\ell}$ from the upper boundary specifically punishes the false claim that the state is h.) Therefore, the best equilibrium will resort to Region 1b behavior only if Region 1a behavior yields insufficient punishment power to meet the sender's incentive constraint. This happens only when the lower boundary is "in the way," that is, when it prevents $w_{H,\ell}$ from being lowered any further.

Nothing in the analytical results ensures that Regions 1 and 2 are intervals. It is not clear that the sender's incentives can be maintained by continuation values on the upper boundary for all $\beta > \underline{\beta}$, so that alternations between Regions 1 and 2 are conceivable, where the continuation value after a dishonest report might have to be dropped below the upper boundary. Interestingly, the numerical computations actually display examples of this, at low discount factors. But for fairly patient senders, the regions have the simplest possible structure:

Numerical Property 3. For $\delta \geq 0.91$, Regions 1, 2 and 3 are all intervals.

Does the provision of incentives ever necessitate the use of the extremely inefficient Region 1b? The numerical computations reveal rare instances of that. But this disappears as the sender displays a reasonable amount of patience:

Numerical Property 4. For $\delta \geq 0.91$, Region 1b is empty, so that the sender's advice is always followed at any point on the upper boundary.

Together with the equilibrium structure, Properties 1, 3 and 4 enormously simplify the picture of equilibrium behavior.

5.4. Lower Boundary Behavior

We make just a few remarks about the lower boundary. It too comprises three regions. Again, if the sender's value $\underline{V}(\beta)$ exceeds the critical value $[\delta + \mu_0(1-2\delta)]\,/\,[1-\delta\mu_0]$, the sender chooses $\pi(\beta) = 1$ and the receiver follows his advice with certainty $\alpha(\beta) = 1$. This is the analogue of the upper boundary's Region 3.

At lower β, $\underline{V}(\beta)$ requires the sender to randomize and two regions emerge, similarly to Section 5.3. In the worst equilibrium, the truthtelling continuation values, $w_{H,h}$ and $w_{L,\ell}$, are on the lower boundary. Ideally, the receiver would ignore the sender's message, $\alpha(\beta) = 0$, so that there would be no chance of H. But this requires $\pi(\beta) = \bar{\pi}(\beta)$ by a variant of Lemma 4. This may or may not be the best way of destroying value, as a larger $\pi(\beta)$ yields a larger $\beta_{L,\ell}$ and hence a larger $w_{L,\ell} = \underline{V}(\beta_{L,\ell})$. The resolution of this tradeoff distinguishes Regions 1 and 2 for the lower boundary.

In Region 1, the receiver completely ignores the sender's message, so $\alpha(\beta) = 0$. This being the case, to make the sender indifferent requires giving him equal continuation payoffs: $w_{H,\ell} = w_{L,\ell}$. Moreover, $\pi = \bar{\pi}$, as already discussed. For some reputations, this is the best configuration to hurt the sender.

We say we are in Region 2 if instead the most punishing configuration is to have both $w_{H,\ell}$ and $w_{L,\ell}$ on the lower boundary, their difference providing the incentive for the sender to be weakly willing to report truthfully. This mirrors Region 2 on the upper boundary, where both rewards in the low state were on the upper boundary. In both these Regions 2 (upper and lower boundaries), $\pi(\beta) < \bar{\pi}(\beta)$ and hence $\alpha(\beta) = 1$.

5.5. Equilibrium Path

Starting from some reputation β, let us follow an equilibrium across periods, for example starting on the upper boundary. Say β is in Region 2. If the state is h, both types of the sender reveal it by Lemma 2, the receiver chooses H by Lemma 4, and no updating occurs. If instead the state is ℓ, the sender randomly recommends H or L and his reputation declines or improves, respectively. Since $\pi(\cdot) > \pi_B$ by Lemma 3, there is an overall downward drift in reputation.[12] Once reputation enters Region 1 and the sender recommends action H in state ℓ, his continuation value leaves the upper boundary. Either this non-maximal continuation value is interior or on the lower boundary. If it is interior, there are many ways to deliver it: one of them is to do a public randomization and, depending on the result, follow either the upper boundary or the lower boundary equilibrium behavior.

5.6. Comparison to Bénabou and Laroque (1992)

Bénabou and Laroque (1992) models a financial journalist who gets imperfect information each period about a traded asset. His type, chosen once and for all, is either

[12]As the model meets the conditions of Cripps, Mailath and Samuelson (2004), the reputation β tends toward 0 in the long run.

a rational, profit-maximizing type who is willing to misinform the market to create a trading opportunity for himself, or compulsively honest (the behavioral type). This means that, as in our model, the sender can report dishonestly without entirely losing his reputation. (Since the signal the sender receives himself is noisy, an erroneous recommendation can be an honest mistake.) Attention is limited to Markov Perfect equilibrium (MPE), where behavior depends only on the sender's current reputation. Notice that the journalist always has a myopic interest in deceiving receivers: whether his private signal is favorable or unfavorable, he would like to create the opposite impression so that he can trade profitably. By contrast, in the class of games we study, the sender wishes to inform the receiver correctly in the high state of the world, and incorrectly in the low state (in our leading example, the health agency always wishes that the public adopts strict health practices). This turns out to be analytically more challenging.

Specifically, think of a function-to-function version of correspondence B from Section 4.1: To any increasing, continuous, continuation payoff function $w : \beta \to [\underline{v}, \bar{v}]$, let b associate another payoff function $b(w) : \beta \to [\underline{v}, \bar{v}]$, which gives the (static Nash) equilibrium payoff of a game whose payoffs are the per-period payoffs augmented with w. Bénabou and Laroque (1992) show that in their model, b is a contraction, and hence, they display a unique MPE in the class of increasing, continuous solutions. In our model, b is not a contraction, and typically, no MPE exists. We study the set of Perfect Bayesian equilibria (PBE), giving special attention to the sender-preferred solutions.

6. Behavioral Convergence

We characterize the sender-preferred equilibrium behavior in the limit as $\delta \to 1$ by establishing an asymptotic equivalence to the commitment solution in the Bayesian persuasion literature. Given the intense interest in static Bayesian persuasion and information design, it is encouraging to see support for its equilibrium behavior in a dynamic setting without commitment. The argument behind behavioral convergence relies on a novel way of evaluating the sender's equilibrium payoffs, which we term "stationary promise-keeping" and explain in 6.2. Behavioral convergence also sheds light on how reputation concerns allow asymptotic efficiency, that is, allow a sender to meet the incentives of a random informational strategy without destroying surplus (see the discussion in Section B.2 of the Online Appendix for further context).

In Region 1, the rate of exploitation, $\pi(\beta)$, is easily shown to be increasing in β, and in Region 3, it has risen to 1. We expected and tried to prove that π is increasing everywhere in Region 2. The numerical computations revealed that for low and moderate discount factors, there are failures of monotonicity of π for some parameter combinations. For all except one peripheral combination,[13] monotonicity of

[13]This failure occurs at $(\mu_0, \pi_B) = (.45, .18)$ for $\delta = .995$ and $\delta = .999$. Here, μ_0 is so high that

π is restored for high values of δ. Whenever π is monotone for a particular parameter pair (μ_0, π_B) at least asymptotically in δ, we can provide strong characterizations of the behavior of patient senders. Accordingly, for all the results of this Section, we assume:

Assumption (Increasing π) For the given (μ_0, π_B) under investigation, there exists $\underline{\delta} \in (0, 1)$ such that for all $\delta \geq \underline{\delta}$, $\pi(\beta)$ is strictly increasing in reputation β.

6.1. Asymptotic Behavior

The sender-preferred equilibrium of the dynamic game converges to the *same* behavior as $\delta \to 1$ for all but a vanishing set of reputations, and that behavior is given by the Bayesian persuasion solution defined in Section 3. A sequence of demanding lemmas (Lemmas 7-13 in the Appendix) supports these behavioral convergence results, reported in Proposition 4.

We first present the central result and discuss its economic implications. Then 6.2 goes back to supply some intuition for the analysis leading to those results.

Proposition 4. (Behavioral Convergence) *For any behavioral type* $\pi_B < \pi^*$ *and* $\epsilon > 0$*, there exists* $\underline{\delta} \in (0, 1)$ *such that*

$$\pi^* \leq \pi(\beta) \leq \pi^* + \epsilon$$

for all $\beta \in [\epsilon, 1 - \epsilon]$ *and* $\delta \geq \underline{\delta}$.

Corollary 1. (Behavioral Convergence) *For any type* $\pi_B < \pi^*$, $\epsilon > 0$ *and number* $T \in \mathbb{N}$*, there exists* $\underline{\delta} \in (0, 1)$ *such that for any* $\delta \geq \underline{\delta}$, $\beta_0 \in [\epsilon, 1-\epsilon]$ *and any realization of* $\{\beta_t\}_{t=1}^T$ *in the (stochastic) equilibrium path,* $|\pi(\beta_t) - \pi^*| < \epsilon$ *and* $\alpha(\beta_t) = 1$ *for all* $t = 0, \ldots, T$.

Proof (of Corollary): Take any $\pi_B < \pi^*$, $\epsilon > 0$ and number T. For any given $\delta \in (0, 1)$, let β_t^H be the reputation value obtained from $\beta_0^H = \epsilon$ after t occurrences of $(\theta, m) = (\ell, H)$. Since $\pi(\beta) \leq \bar{\pi}(\beta)$ for all $\beta \in (0, 1)$,

$$\underline{\beta}_\epsilon \equiv \lambda_0^T \epsilon \leq \lambda_0^t \beta_0^H \leq \beta_t^H \quad \text{for all } t = 0, \ldots, T.$$

Similarly, let β_t^L be the reputation value obtained from $\beta_0^L = 1 - \epsilon$ after t occurrences of $(\theta, m) = (\ell, L)$. Temporarily, assume that $\pi(\beta_t^L) \leq \pi^* + \epsilon$ for all $t = 0, \ldots, T - 1$. Let $\bar{\beta}_\epsilon$ be the solution of

$$\frac{1 - \bar{\beta}_\epsilon}{\bar{\beta}_\epsilon} = \frac{\epsilon}{1 - \epsilon} \left[\frac{1 - \pi^* - \epsilon}{\pi_B} \right]^T.$$

there is only the mildest of incentive problems. The failure may be a computational artifact: for reputations high in Region 2, rewards for truthtelling are in Region 3, where $\bar{V}$ is staggeringly steep. (For the parameters in question, at some points, the slope of $\bar{V}$ exceeds 10^{40}.) This tends to make the numerical determination of π unstable.

Then

$$\frac{1-\beta_t^L}{\beta_t^L} \geq \frac{1-\beta_0^L}{\beta_0^L}\left[\frac{1-\pi^*-\epsilon}{\pi_B}\right]^t \geq \frac{1-\bar{\beta}_\epsilon}{\bar{\beta}_\epsilon},$$

or equivalently, $\beta_t^L \leq \bar{\beta}_\epsilon$ for all $t = 0, \dots, T$. Since $\lambda_0 < 1$ and $\pi_B < \pi^* < \pi^* + \epsilon$, $\underline{\beta}_\epsilon < \epsilon$ and $\bar{\beta}_\epsilon > 1 - \epsilon$.

Finally let $\bar{\epsilon} = \bar{\pi}(\underline{\beta}_\epsilon) - \pi^*$. Since $\bar{\pi}(0) = \pi^*$ and $\bar{\pi}$ is a strictly increasing function, $\bar{\epsilon} > 0$. Pick any $0 < \hat{\epsilon} < \min\{\underline{\beta}_\epsilon, 1 - \bar{\beta}_\epsilon, \bar{\epsilon}\}$. By Proposition 4, there exists $\underline{\delta} \in (0,1)$ such that $\pi^* \leq \pi(\beta) \leq \pi^* + \hat{\epsilon} < \pi^* + \epsilon$ for all $\beta \in [\hat{\epsilon}, 1-\hat{\epsilon}]$ and $\delta \geq \underline{\delta}$. Fix $\delta \geq \underline{\delta}$. Then, for all $\beta \in [\underline{\beta}_\epsilon, 1]$, $\bar{\pi}(\beta) \geq \pi^* + \bar{\epsilon} > \pi^* + \hat{\epsilon} \geq \pi(\beta)$, which implies that $\alpha(\beta) = 1$. Since $1 - \hat{\epsilon} > 1 - \bar{\beta}_\epsilon$, the assumption made earlier is satisfied and indeed $\beta_t^L \leq \bar{\beta}_\epsilon$ for all $t = 0, \dots, T$. Finally, since $\pi(\beta)$ is an increasing function in $[\hat{\epsilon}, 1-\hat{\epsilon}] \supset [\underline{\beta}_\epsilon, \bar{\beta}_\epsilon]$, for any $\underline{\beta}_\epsilon < \beta < \beta' < \bar{\beta}_\epsilon$ we have that $\beta_{H,\ell} < \beta'_{H,\ell}$ and $\beta_{L,\ell} < \beta'_{L,\ell}$. Therefore, for any $\beta_0 \in [\epsilon, 1-\epsilon]$, $\beta_t \in [\underline{\beta}_\epsilon, \bar{\beta}_\epsilon] \subset [\hat{\epsilon}, 1-\hat{\epsilon}]$ for all $t = 1, \dots, T$, and thus $\pi^* < \pi(\beta_t) < \pi^* + \epsilon$ and $\alpha(\beta) = 1$ for all $t = 0, \dots, T$. □

We make several remarks. First, convergence in behavior is not an immediate consequence of convergence in value (Proposition 1 and FL).[14] In the dynamic model, behavior can change substantially while payoffs remain constant. This is because of two features of behavior that are substitutes in value production: a large $\pi(\cdot)$ increases value by inducing more H, but a steep $\pi(\cdot)$ destroys value through adverse reputational dynamics (see Lemma 10 in the Appendix). At first glance, then, it would seem possible to sustain the roughly constant value of $\overline{V}$ that one sees in FL by having $\pi(\cdot)$ increase at the right rate. Proposition 4 shows, however, that any substantial increase in π leads to a feeding frenzy: more and more explosive increases in π are needed at successively higher levels of β, which is unsustainable.

Furthermore, Proposition 4 and its corollary hold for all $\pi_B < \pi^*$, as illustrated in Figure 3. Whereas convergence in value requires π_B to be close to π^*, the analogous result for behavior does not: even when π_B is much lower than π^*, behavior in the sender-preferred PBE converges. But it converges to π^*, not to π_B! And the corresponding value converges to neither the average payoff associated with always being trusted and cheating according to π^* (which is the Bayesian persuasion value v^*) nor the average payoff associated with always being trusted and cheating according to π_B (which was defined as v^{π_B} in Section 3). Let us look more closely at the difference between the case where π_B is close to π^* and the case when it is not.

In either case, because $\pi(\cdot) > \pi_B$ by Lemma 3, reputation drifts inexorably downward and tends to 0, where the sender's per period payoff is no more than μ_0. This long-run scenario is consistent with Cripps, Mailath and Samuelson (2004). In determining the sender's average discounted value, there is a race between how patient

[14]In contrast, behavioral convergence follows immediately from payoff convergence in the standard repeated game: if $\overline{V}(0;\delta) = v^*$, then the sender must get v^* in every period, which requires playing the Bayesian persuasion solution in every period.

he is, which gives a lot of weight to vanishing reputations, and how long it takes to lose reputation, which is an updating phenomenon. The latter effect wins when π_B is virtually π^*, because the lying frequency of the behavioral and rational types are almost identical even in Region 1, so the downward reputational movement toward low payoffs is overwhelmingly slow. When instead π_B is much less than π^*, the behavioral and rational types have quite distinct lying probabilities in Region 1 (and in Region 2), hence the eventual low payoffs are not so far away, and have a noticeable impact on the weighted average payoffs. This is why, in this case, the resulting average discounted payoff is less than the Bayesian persuasion value, even though behavior throughout Region 2 is close to the Bayesian persuasion ideal.

Proposition 4 has further dramatic consequences: Regions 1 and 3 vanish asymptotically. If Region 1 did not vanish, then $\pi(\cdot)$ would be equal to $\bar{\pi}(\cdot)$ for reputations β bounded away from 0 and hence grow to values bounded away from π^*. This would contradict Proposition 4, which asserts that $\pi(\cdot)$ gets arbitrarily close to π^*. Moreover, $\pi(\beta)$ can equal 1 only if $\beta > 1-\epsilon$, hence Region 3 also vanishes asymptotically, leaving Region 2 and its efficient regime to fill the entire space.

6.2. Intuition and Stationary Promise-Keeping

Flatness. We first argue that for any $\pi_B < \pi^*$, $\overline{V}$ becomes almost flat over the entire reputation space. Think of the incentive needed to keep the sender indifferent in Region 1 or 2:

$$\overline{V}(\beta_{L,\ell},\delta) - \overline{V}(\beta_{H,\ell},\delta) \leq \frac{1-\delta}{\delta}.$$

The right side vanishes as $\delta \to 1$, because the average discounted value of an extra action H becomes negligible in the long run. Therefore, the vertical step sizes become miniscule while the horizontal step sizes $\beta_{L,\ell} - \beta_{H,\ell}$ are bounded away from zero, except for β extremely close to 1 or 0, where a new message causes little updating. This makes for a *very* flat $\overline{V}$ function. The only other way for $\beta_{L,\ell} - \beta_{H,\ell}$ to vanish would be for there to be almost no information about the sender's type in his message, that is, for π to be arbitrarily close to π_B. But this is impossible because π starts at π^* and is weakly increasing, so $\pi - \pi_B$ is uniformly bounded below.

Lemma 11. *For any $\epsilon > 0$, there exists $\underline{\delta} \in (0,1)$ such that $\overline{V}(1-\epsilon,\delta) - \overline{V}(\epsilon,\delta) \leq \epsilon$ for all $\delta \geq \underline{\delta}$.*

Stationary Promise-Keeping. When $\delta \to 1$, $\overline{V}(\beta,\delta)$ converges to the same value for all β, a value which depends on π_B: what is that value? To answer, we propose a novel way of evaluating the sender's expected payoff at reputation β, called stationary promise keeping. In a sender-preferred equilibrium starting in Region 1 or 2, if the rational sender observes ℓ, both messages H and L are best responses for him. Hence, his expected utility can equivalently be evaluated by right or left promise keeping, defined in 5.3.2 above. These are value decompositions that break the payoff into a weighted average of today's return and the appropriate continuation values. If little is known about the latter, this does not say much about the average value. One could iterate left promise-keeping many times, progressively unpacking the continuation values, but this leads through a nonlinearly changing environment and is hard to evaluate.

Instead, starting at some β_0 in Region 2, the sender could maintain his reputation as steady as possible, always saying L if current reputation is strictly below β_0 and $\theta = \ell$, and always saying H otherwise. This "reputation maintenance" strategy keeps his reputation permanently within the interval $[\beta_H, \beta_L]$, where $\pi_0 = \pi(\beta_0)$, $\beta_H = \beta_{H,\ell}(\beta_0, \pi_0)$ and $\beta_L = \beta_{L,\ell}(\beta_0, \pi_0)$. As long as $[\beta_0, \beta_L]$ is included in Region 2, both messages are optimal, the continuation values remain on the upper boundary, and hence this reputation maintenance strategy has the same payoff as his equilibrium strategy.[15]

The idea now is to approximate this value by approximating the flow rate of actions H it can induce, which is what the sender cares about if he is patient. In the case where π is flat across the interval $[\beta_H, \beta_L]$, it is remarkably easy to find the proportion of H that is possible. Forget about following the details of the path that the sender's reputation will follow when he engages in reputation maintenance, and focus instead on long run likelihoods. The receiver entertains two hypotheses: she is watching a time series generated either by a behavioral type who uses π_B, or a rational type who uses π. Her posterior will be unaffected if the number of H messages makes the observed sequence of messages as likely under the one hypothesis as under the other. Lemma 9 in the Appendix does that computation, which is approximately

$$V^M(\pi_0, \pi_B) = \mu_0 + (1 - \mu_0) \frac{\log\left[\frac{1-\pi_B}{1-\pi_0}\right]}{\log\left[\frac{\pi_0(1-\pi_B)}{(1-\pi_0)\pi_B}\right]} \qquad \text{for} \quad \pi_0 \in [\pi^*, 1).$$

At very low reputations β, $\pi(\beta) = \bar{\pi}(\beta)$ is close to π^*. Thus, for δ high enough that $\overline{V}$ is flat around β, stationary promise-keeping suggests that (if reputation were maintained within Region 2) $\overline{V}(\beta, \delta)$ should be close to $V^M(\pi^*, \pi_B)$.

Behavior. Proposition 4 above and its Corollary assert convergence, as $\delta \to 1$, of the sender-preferred equilibrium behavior to values that depend neither on β nor π_B. Moreover, those values coincide with π^* and $\alpha = 1$, the solution to the static Bayesian persuasion problem with commitment. This offers an alternative interpretation of much current work on information design, an interpretation relying on reputational dynamics rather than on commitment.

[15] Recall that the standard devices of left and right promise-keeping correctly evaluate the sender's equilibrium continuation value, even though neither conforms to equilibrium behavior. The same is true of stationary promise-keeping: holding the receiver's strategy fixed at its equilibrium specification, anything in the support of the sender's equilibrium distribution yields the correct equilibrium continuation value.

7. Multiple Behavioral Types

In this Section, we allow a more general treatment with many behavioral types and make use of our algorithm to demonstrate an intriguing point about whom the sender should claim to be. We show first by numerical example and then by a limit theorem, that the sender may put higher weight on a lower π_B (further from the Bayesian persuasion ideal) than on a higher one.

In this alternative treatment, finitely many *transparent* types $\Pi_B = \{\pi_B^1, \ldots, \pi_B^K\}$ (each representing some probability of lying in state ℓ) are available[16] and, for expositional ease, equally likely a priori, and the sender first announces which one of them he is. Since, following that announcement, receivers only need to form a belief about whether the sender is rational or the type he announced, the analysis from 5.1 to 5.5 can be interpreted as applying to a subgame in which that announcement has already been made.

Assumption. $\pi_B^n < \pi^*$ for all $n = 1, \ldots, K$.

In virtue of this assumption, all types in Π_B are trusted by the receiver.[17] After the sender's announcement, receivers form a belief about whether the sender is rational or the type he announced. For the sake of argument, assume the sender gets his best PBE following his declaration of type.

We make several observations about equilibrium behavior in this game. First, the starting reputation in a subgame *decreases* as the probability that the rational sender announces that type increases. Indeed, if the rational sender announces π_B^n with a very small probability, then a receiver in that subgame will believe that he is facing the true π_B^n with high probability.

Second, the rational sender must announce each type in Π_B with strictly positive probability in equilibrium, for otherwise announcing a type that is never reported would win a reputation of 1 (to be that type), which would be a lucrative deviation. This implies that the rational sender must be indifferent over all behavioral types.

Third, the previous point implies that if the sender does better in the subgame following π_B^n than in the one following π_B^k (if, counterfactually, he started each subgame at the same β), then he will announce π_B^n with *higher probability* than π_B^k, in order to induce different starting reputations which equalize the two expected payoffs.

Finally, one might think it better for the sender to claim to be a better type, that is, one closer to π^* (since it is trusted anyway). The naive intuition for this is that one should get more actions H by passing for a higher π^*. In the rest of this Section, we show that things are not so simple away from the limit $\delta \to 1$, starting with the following numerical observation:

[16] In the reputational literature (see for example Abreu and Pearce (2007)), transparency is a convenient assumption according to which the sender announces a type at the beginning of the game and a behavioral sender is assumed to announce his type honestly.

[17] One can show that a behavioral type $\pi_B \geq \pi^*$ is not helpful to the sender.

Numerical Property 6. *Both for* $\delta = .9$ *and for* $\delta = .95$*, the rational sender does strictly better in equilibrium in the subgame following* $\pi_B = 0.2$ *than in the subgame following* $\pi_B = 0.3$*,* $\bar{V}_{0.2}(\beta, \delta) > \bar{V}_{0.3}(\beta, \delta)$*, for all reputations* $\beta \in [0, 0.85]$*.*

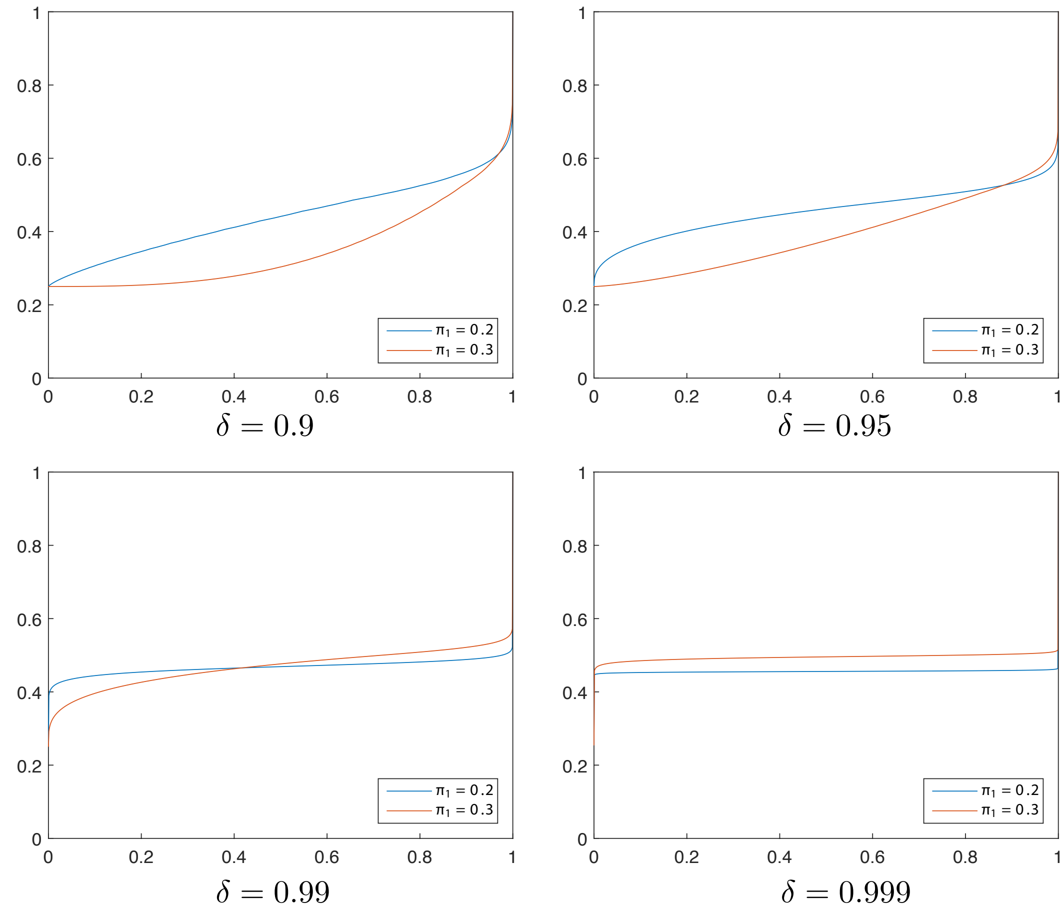

Figure 4: $\overline{V}$ for $\pi_B = 0.2$ and $\pi_B = 0.3$

Although the rational sender is indifferent about which type to announce in equilibrium, it is quite remarkable that a worse type can be more favorable to the sender at most reputations, and hence announced with a larger probability. The argument that the sender "should" induce more H by passing for a higher type neglects the receiver's equilibrium response to the presence of a worse (from her perspective) behavioral sender. This manifests itself in $\bar{\pi}$, which is decreasing in π_B, so that the sender's rate of exploitation decreases with π_B. Since the sender's present discounted payoff can be measured via a streak of L recommendations, only interrupted by an H when in Region 3 (this right promise-keeping strategy is weakly optimal), the sender's payoffs are the average discounted number of entries or re-entries into Region 3. Since $\bar{\pi}$ and π are often larger at $\pi_B = 0.2$ than at $\pi_B = 0.3$, the Bayesian step $\beta_{L,\ell} - \beta$ at any given reputation is also larger at $\pi_B = 0.2$ than at $\pi_B = 0.3$ (because messages discriminate better about sender type). Thus, it is possible that the sender travels faster to Region 3 and accrues more actions H.

This result is illustrated in Figure 4. When δ is close to 1, things go as expected: the sender does better if π_B is closer to π^*, as he earns almost his Bayesian persuasion value except for extreme values of β. But at more modest levels of δ, such as .9 and .95, the sender does better for a wide range of reputations in the subgame where he has a relatively low π_B to imitate, far from the Bayesian persuasion ideal of π^*.

There is a more general way to see the danger of π_B close to π^*. Fix any parameters μ_0 and δ. The vertical step size (the reward for reporting L truthfully instead of lying at some reputation β in Region 2) is fixed at $(1-\delta)/\delta$, which limits the number of (horizontal) Bayesian steps $\beta_{L,\ell} - \beta_{H,\ell}$ across Region 2. But by choosing π_B close enough to π^*, we can make the horizontal step sizes vanishingly small (the behavioral and the rational types behaviors are almost indistinguishable). Since the number of vertical steps across Region 2 is bounded, so is the number of horizontal steps. Lemma 14 shows that one can therefore choose a reputation β close to 1 at which there is a negligible difference between $\bar{V}(\beta)$ and $\bar{V}(\beta_{L,\ell})$. Right promising-keeping at that β shows $\bar{V}(\beta)$ to be little more than $\bar{V}(0)$. In other words, for any δ, no matter how high, a sufficiently high π_B has disastrous value consequences. Lemma 14 also shows the further consequence that Region 1 would then extend across almost the entire reputation space.

Lemma 14. *For any fixed $\delta \in (0,1)$ and for any $0 < \epsilon \leq (1-\delta)/[\delta(1-\mu_0)]$, there exists $\bar{\gamma} > 0$ such that for all $\pi_B \in [\pi^* - \bar{\gamma}, \pi^*)$, $\overline{V}(\beta) \leq \mu_0 + \epsilon$ for all $\beta \in [0, 1-\epsilon]$. Moreover, if Region 1b is empty then Region 1 contains $[0, 1-\epsilon]$.*

8. Related Literature

Crawford and Sobel (1982) and Kamenica and Gentzkow (2011) are the seminal contributions on Cheap Talk and Bayesian persuasion (or information design), respectively. These two literatures differ in the sender's ability to commit to a communication protocol. Our work connects them via implicit enforcement. In an earlier related result, Aumann and Maschler (1995) studied optimal disclosure in infinitely repeated zero-sum games between two players maximizing their long-run average payoffs, when only one of them is informed about the state. More recent works include Hörner, Takahashi, and Vieille (2015), Ely (2017), Best and Quigley (2017), and Margaria and Smolin (2018). Ely (2017) studies dynamic persuasion mechanisms (with commitment) in a model in which a long-lived sender observes the evolution of a stochastic process and communicates with short-lived receivers. Our paper shares with Best and Quigley (2017) the goal of understanding how repetition can substitute for the commitment assumption in information design. They look for ways such as "review aggregation" to change the informational conditions and escape the negative implications of what we document in Proposition A.2 in the Online Appendix. A well-calibrated review aggregator that creates long delays in reporting on the average veracity of a sender's messages, while keeping receivers (but not the sender) unclear about the timing of the next review, can approximate the sender's payoff with commitment. By contrast, our approach is to accept the existing informational limitations and explore how reputational mechanisms can overcome those limitations.

Our paper is also related to the literature on repeated Cheap Talk with reputation, such as Sobel (1985), Bénabou and Laroque (1992) and Morris (2001). In these papers, a privately informed and long-lived sender interact over a finite horizon with a myopic receiver. In Sobel (1985), the sender can be a good type that always speaks the truth or a bad type whose preferences are severely misaligned with the receiver's. The bad type first builds a reputation for being a friend and times his deceit for maximal gains. Bénabou and Laroque (1992) analyze a version of Sobel's game in which the sender has noisy information. They limit attention to MPEs. Because of the different alignment of interests between sender and receiver in our model (sometimes aligned, sometimes misaligned, as in KG), there are usually no Markov perfect equilibria in the class we study. We analyze the entire set of PBEs, and then focus on the sender-preferred ones. In Morris (2001), the good type has the same preferences as the receiver, while the bad type wants the receiver to always choose the same action (independently of information). In this model, the good type's effort to distinguish herself from the bad type results in no information being conveyed in equilibrium. This has echoes in other papers such as Ely and Valimäki (2003) and Mailath and Samuelson (2001), where the dynamics are driven by the same desire for a good type to separate itself from a bad type—and failing to do so in equilibrium. In Mailath and Samuelson (2001), reputation regains a role if types can change over time (see also Phelan (2006) for an application in political economy). Our work is closer to Sobel (1985), in the sense that our good type is committed to a relatively trustworthy behavior;[18] but given that behavior involves randomization and the time horizon is infinite, the dynamics are much richer. Finally, Ottaviani and Sorensen (2006a,b) study a single round of Cheap Talk interaction. Nonetheless, the sender may want to appear more precisely informed than he is for reputational reasons, presumably to enjoy higher payoffs in the future.

There is also a literature on multi-round Cheap Talk, with "long-run" players involved in dynamic communication, sometimes bilateral. The state is usually drawn once and for all at the beginning. In Aumann and Hart (2003), two players (one informed and one uninformed) play a finite normal form game. They exchange (possibly, infinitely many) messages, before simultaneously choosing actions. In contrast, in Golosov, Skreta, Tsyvinski and Wilson (2014), only the informed party sends messages and the uninformed party chooses actions. Krishna and Morgan (2004) add a long communication protocol to Crawford and Sobel (1982) and show that it leads to Pareto-improving information transmission. Goltsman, Hörner, Pavlov and Squintani (2009) characterize and compare such optimal protocols. Forges and Koessler

[18]We use the standard reputation framework from Kreps and Wilson (1982), Milgrom and Roberts (1982), and Kreps, Milgrom, Roberts and Wilson (1982), Fudenberg and Levine (1989, 1992).

(2008a,b) allow for a long protocol in a setup where messages are certifiable. Pei (2021) considers a patient sender who has a cost of lying that is private information. He gives conditions under which the sender can attain the Bayesian persuasion payoff.

Finally, computational approaches to complex dynamical systems have been used to pave the way for theoretical results in quantitative dynamic public finance (see Golosov et al. (2017) for a survey) and asset pricing (e.g., Borovicka and Stachurski (2020)), just to mention some examples. In the former, a model with distortions would typically be simulated numerically (for instance, when the utility function is nonseparable) and then theoretical arguments would guide policymaking. In Borovicka and Stachurski (2020), existence and uniqueness of equilibrium asset prices in infinite-horizon models rely on necessary and sufficient conditions that cannot be established theoretically but that hold numerically. Phelan and Stacchetti (2001) study a Ramsey tax model, modifying the APS algorithm to compute its equilibrium value correspondence. In general dynamic incentives problems, Renner and Scheidegger (2020) develop APS-based computational methods for computing solutions that standard techniques cannot derive analytically, such as extensions of Fernandes and Phelan (2000).

9. Conclusion

A long-lived sender can build, maintain or run down his reputation for a degree of honesty in reporting information that arrives period by period. If he is extremely patient, a powerful result of Fudenberg and Levine (1992) guarantees him an average discounted payoff almost as high as the Bayesian persuasion value of Kamenica and Gentzkow (2011) for a sender who can commit himself to a particular random information protocol. Here, we investigate theoretically and computationally what kind of equilibrium behavior supports this value result. In addition, we study both value and behavior for discount factors distant from 1.

For general discount factors, sender-preferred equilibria have a three-region structure. Usually each of these regions is an interval in reputation space. In Region 3, where reputation is very high, the receiver always acts on the senders advice, although the (rational) sender always claims the state is high even when it is low. He is "cashing in" on his high reputation. In Region 2, where reputation is more moderate, the receiver again trusts the sender, who randomizes between reporting honestly or dishonestly. The randomization probability at any reputation level is constructed such that lying is punished just enough to keep the sender indifferent between reporting the truth or lying. Whichever he does, at the new prevailing reputation, the continuation value is maximal in the set of values at that reputation, so we say that Region 2 involves efficient provision of incentives. By contrast, at the lower reputation levels of Region 1, the sender's continuation value after a lie is not on the upper boundary of the value correspondence. Here, there is inefficient provision of incentives.

Although behavior is always as described above, the computational results reveal two interesting phenomena, both of which disappear if the sender is fairly patient.

First, Regions 1 and 2 might not be intervals. For extremely low reputations, Region 1 behavior always applies. But for some parameter combinations, as reputation is increased, one passes into Region 2 and then back into 1. Although behavior within each region is preserved, the regions themselves do not have the extremely simple structure that asserts itself for higher discount factors. Secondly, in Region 1, lowering the continuation value from the upper boundary after a lie may not be enough to deter lying, because the lower boundary of the value correspondence is "in the way": it is insufficiently harsh. In that case, the receiver sometimes ignores the sender's advice, reducing his incentive to lie. This is a further inefficiency of Region 1: sometimes the truth is disregarded by the receiver, hurting both sender and receiver.

Section 7 offers some insight into what happens if the sender can choose from a set of behavioral types to imitate, rather than there being only one type to imitate. A natural guess here would be that types closer to the ideal Bayesian persuasion commitment probability should be imitated more often in equilibrium. This is incorrect. If a type very close to that commitment type were chosen frequently, the sender would receive little more than his truthtelling value, which is the payoff in the standard repeated game, in which he has no behavioral types to imitate at all. It turns out that, just as the tailor has to cut the cloth to suit the purse, a sender of a particular degree of patience cannot afford to imitate frequently a very high behavioral type.

None of the preceding results requires the discount factor δ to approach 1. We were surprised by what happens when that limit exercise is performed. Section 6 undertakes the asymptotic analysis, in the presence of a single behavioral type. Recall our notation π_B for the probability with which the behavioral type lies when the state is low. This is a lower number that the Bayesian persuasion probability (of lying in the low state) π^* (it is easy to show that otherwise, the presence of that behavioral type serves no purpose), and consequently, as δ approaches 1, the sender's payoff does not approach the Bayesian persuasion value. In spite of that, in regular cases,[19] the sender's equilibrium *behavior* does approach the Bayesian persuasion probability! At each fixed reputation value in $(0, 1)$, asymptotically the sender behaves as though he had committed himself as in KG. For any particular δ, this does not yield the Bayesian persuasion value, because π^* exceeds π_B, and in the long run, the sender loses his reputation and receives low payoffs. But a novel technique for calculating equilibrium payoffs, which we call *stationary promise-keeping*, easily computes the true limiting average discounted value. If one chooses π_B close to π^* and then chooses δ sufficiently high, both behavior and value will closely resemble the Bayesian persuasion solution of KG, providing a dynamic interpretation of Bayesian persuasion analysis without resorting to any legal commitment.

[19]Here we assume that for sufficiently large δ, equilibrium lying probabilities increase with reputation. The computational results suggest that this is usually the case.

Appendix

Proof of Lemma 1: Clearly $E(\pi, \alpha, w) \in [\underline{v}, \overline{v}]$ for each $\beta \in \Delta(\hat{\Pi})$ and each tuple (π, α, w) admissible for W at β. Therefore

$$B(W) \subseteq \Delta(\hat{\Pi}) \times [\underline{v}, \overline{v}]$$

is a bounded set. Let $\{(\beta^k, z^k)\} \subseteq B(W)$ be a sequence such that $\beta^k \to \beta$ and $z^k \to z$. We now show that $(\beta, z) \in B(W)$. For each $k \in \mathbb{N}$, there is a tuple (π^k, α^k, w^k) admissible for W at β^k such that $z^k = E(\pi^k, \alpha^k, w^k)$. Since $\Delta(M)^\Theta$, $\Delta(A)^M$ and $[\underline{v}, \overline{v}]^{M\times\Theta}$ are all compact sets, without loss of generality we can assume that $\pi^k \to \pi^\infty$, $\alpha^k \to \alpha^\infty$ and $w^k \to w^\infty$ for some $\pi^\infty \in \Delta(M)^\Theta$, $\alpha^\infty \in \Delta(A)^M$ and $w^\infty \in [\underline{v}, \overline{v}]^{M\times\Theta}$. One can check that $(\pi^\infty, \alpha^\infty, w^\infty)$ is admissible for W at β. Finally,

$$z = \lim_{k\to\infty} z^k = \lim_{k\to\infty} E(\pi^k, \alpha^k, w^k) = E(\pi^\infty, \alpha^\infty, w^\infty).$$

Therefore, $(\beta, z) \in B(W)$. This establishes that $B(W) \subseteq \Delta(\hat{\Pi}) \times [\underline{v}, \overline{v}]$ is a closed set and hence that $B(W) : \Delta(\hat{\Pi}) \to [\underline{v}, \overline{v}]$ is an upper hemicontinuous correspondence. Since $\tilde{B}(W)(\beta) = \mathrm{co}(B(W(\beta)))$ for each $\beta \in \Delta(\hat{\Pi})$, $\tilde{B}(W)(\beta)$ is a compact convex set for each $\beta \in \Delta(\hat{\Pi})$ and $\tilde{B}(W) : \Delta(\hat{\Pi}) \to [\underline{v}, \overline{v}]$ is also an upper hemicontinuous correspondence.

It remains to show that $B(W)(\beta) \neq \emptyset$ for each $\beta \in \Delta(\hat{\Pi})$. Fix $\beta \in \Delta(\hat{\Pi})$ and consider a simultaneous moves auxiliary game with three players: the sender, the receiver and a "dummy player". The sender's payoff function in the auxiliary game is $E(\pi, \alpha, w)$, and his best reply correspondence is $\mathrm{BR}^S(\alpha, w)$. The receiver's payoff is $u(\theta, a)$, as in the component game (and does not depend on the dummy's actions), and his best reply correspondence is $\mathrm{BR}^R(\beta, \pi)$. We do not specify a payoff function for the dummy player. Instead, we specify directly his "best-reply correspondence":

$$\mathrm{BR}^D(\pi, \alpha) = \prod_{(m,\theta)\in M\times\Theta} W(\beta_{m,\theta}(\cdot|\beta, \pi)).$$

Clearly, for each (m, θ), the function $\pi \to \beta_{m,\theta}(\cdot|\beta, \pi)$ from $\Delta(M)^\Theta$ to $\Delta(\Pi_0)$ is continuous. Since $W \in \mathcal{C}$, the correspondence $\pi \to W(\beta_{m,\theta}(\cdot|\beta, \pi))$ is upper hemicontinuous with convex and compact values.

As remarked earlier, $\mathrm{BR}^S(\alpha, w)$ and $\mathrm{BR}^R(\beta, \pi)$ are convex and compact sets. By the Maximum Theorem, one can also easily check that BR^S and BR^R are upper hemicontinuous correspondences. Let

$$\mathrm{BR}(\pi, \alpha, w) = \mathrm{BR}^S(\alpha, w) \times \mathrm{BR}^R(\beta, \pi) \times \mathrm{BR}^D(\pi, \alpha).$$

It is easy to see that (π, α, w) is an admissible tuple for W at β if and only if (π, α, w) is a fixed point for BR. By Kakutani's theorem, the correspondence BR from $\Delta(M)^\Theta \times$

$\Delta(A)^M \times [\underline{v},\overline{v}]^{M\times\Theta}$ into itself has a fixed point (π,α,w). Therefore, $E(\pi,\alpha,w) \in B(W)(\beta)$ and $B(W)(\beta) \neq \emptyset$. □

Proof of Proposition 3: By definition,

$$W^\infty(\beta) = \lim_{k\to\infty} W^k(\beta) = \bigcap_{k\in\mathbb{N}} W^k(\beta).$$

The intersection of compact and convex sets is a compact convex set, and by the finite intersection property $W^\infty(\beta)$ is non-empty. Thus $W^\infty \in \mathcal{C}$.

We now show that W^∞ is self-generating. Let $\beta \in \Delta(\hat{\Pi})$ and $w \in W^\infty(\beta)$. We need to show that $w \in \tilde{B}(W^\infty)(\beta)$. Since $w \in W^\infty(\beta)$, $w \in W^{k+1}(\beta)$ for all k. Therefore, for each $k \in \mathbb{N}$ there exist $x^k \in [0,1]$ and two tuples $(\pi^{1k},\alpha^{1k},w^{1k})$ and $(\pi^{2k},\alpha^{2k},w^{2k})$ admissible for W^k at β such that $w = x^k E(\pi^{1k},\alpha^{1k},w^{1k}) + (1-x^k)E(\pi^{2k},\alpha^{2k},w^{2k})$. Again, without loss of generality we can assume that $x^k \to x^\infty$, $\pi^{jk} \to \pi^{j\infty}$, $\alpha^{jk} \to \alpha^{j\infty}$ and $w^{jk} \to w^{j\infty}$ for some $\pi^{j\infty} \in \Delta(M)^\Theta$, $\alpha^{j\infty} \in \Delta(A)^M$ and $w^{j\infty} \in [\underline{v},\overline{v}]^{M\times\Theta}$. It is easy to check that $w^{j\infty}_{m,\theta} \in W^\infty(\beta_{m,\theta}(\cdot|\beta,\pi^{j\infty}))$ for each $(m,\theta) \in M\times\Theta$, and hence that $(\pi^{j\infty},\alpha^{j\infty},w^{j\infty})$ is an admissible tuple for W^∞ at β, $j=1,2$. Moreover,

$$\begin{aligned} w &= \lim_{k\to\infty} x^k E(\pi^{1k},\alpha^{1k},w^{1k}) + (1-x^k)E(\pi^{2k},\alpha^{2k},w^{2k}) \\ &= x^\infty E(\pi^{1\infty},\alpha^{1\infty},w^{1\infty}) + (1-x^\infty)E(\pi^{2\infty},\alpha^{2\infty},w^{2\infty}). \end{aligned}$$

Therefore $w \in \tilde{B}(W^\infty)(\beta)$ as was to be shown.

By Proposition 2, $W^\infty \subseteq V$. Conversely, since $\tilde{B}$ is monotone and $V \subseteq \tilde{B}(W^0) \subseteq W^0$, we have $V \subseteq W^k$ for all $k \geq 0$. Therefore, we also have that $V \subseteq W^\infty$. □

Proof of Lemma 2: By contradiction, assume that $\pi(H|h,\beta) < 1$. Then, in equilibrium, it is optimal for the sender to recommend L when $\theta = h$. But (L,h) reveals rationality (since the behavioral type never recommends L when $\theta = h$). Hence $\beta_{L,h} = 0$ and the best continuation value that can be offered in this case is $w_{L,h} = \overline{V}(0) = \mu_0$. Therefore, the sender's expected continuation value after observing $\theta = h$ is bounded above by $\delta\mu_0$. But, keeping $(\pi(H|\ell,\beta),\alpha(\beta),w_{L,\ell},w_{H,\ell})$, make $\pi(H|h,\beta) = 1$ instead. Then $\beta_{H,h} = \beta$ and we can set $w_{H,h} = \overline{V}(\beta) > \mu_0$ (and $w_{L,h} = 0$). The modified admissible tuple is incentive compatible and the sender's expected continuation value after observing $\theta = h$ is $\alpha(\beta)(1-\delta) + \delta\overline{V}(\beta) > \delta\mu_0$. Thus the modified admissible tuple has a strictly higher total value, which is a contradiction. □

Proof of Lemma 3: If $\pi(\beta) \leq \pi_B < \pi^*$, then it is optimal for the sender to recommend L after $\theta = \ell$, and $\beta_{L,\ell} \leq \beta$. Hence $v_\ell \leq \delta\overline{V}(\beta_{L,\ell}) \leq \delta\overline{V}(\beta)$. Meanwhile $v_h \leq \alpha(\beta)(1-\delta) + \delta\overline{V}(\beta) \leq (1-\delta) + \delta\overline{V}(\beta)$. Thus,

$$\overline{V}(\beta) \leq \mu_0[(1-\delta) + \delta\overline{V}(\beta)] + (1-\mu_0)\delta\overline{V}(\beta) \quad \Longrightarrow \quad \overline{V}(\beta) \leq \mu_0,$$

which is not possible since $\overline{V}(\beta) > \mu_0$ for all $\beta > 0$. □

Proof of Lemma 4: By way of contradiction, suppose that $\pi(\beta) > \bar{\pi}(\beta)$ for some β. By definition of $\bar{\pi}$, this implies that the receiver does not follow H recommendations, that is, $\alpha(\beta) = 0$. Since $\pi(\beta) > 0$, H is optimal for the sender at β when $\theta = \ell$, and by Lemma 2, H is also optimal when $\theta = h$. Therefore

$$\overline{V}(\beta) = \mu_0 \delta w_{H,h} + (1 - \mu_0)\delta w_{H,\ell}.$$

Since $w_{H,h} \leq \overline{V}(\beta)$ and $w_{H,\ell} \leq \overline{V}(\beta_{H,\ell})$, we have that $\overline{V}(\beta) \leq (1-\mu_0)\delta\overline{V}(\beta_{H,\ell})/[1 - \mu_0\delta] < \overline{V}(\beta_{H,\ell})$. But $\pi(\beta) > \bar{\pi}(\beta) > \pi_B$ implies $\beta_{H,\ell} < \beta$ and $\overline{V}(\beta_{H,\ell}) < \overline{V}(\beta)$, so the previous inequality implies $\overline{V}(\beta) < \overline{V}(\beta)$, a contradiction. □

Let β_d be such that $\bar{\pi}(\beta_d) = 1$. That is, $\mu(1, \beta_d) = 1/2$. For reputations $\beta > \beta_d$, the receiver strictly prefers to follow the recommendation H even when the sender recommends H for sure in every state.[20]

Suppose for some $\beta > \beta_d$, $\beta_{H,\ell} = \beta_{H,\ell}(\beta, 1)$ is such that $(1-\delta) + \delta\overline{V}(\beta_{H,\ell}) > \delta$, or equivalently such that $\overline{V}(\beta_{H,\ell}) > 2 - 1/\delta$. That is, suppose that when the sender observes $\theta = \ell$, the total value of recommending H (today) exceeds the value of getting 0 today and 1 in every period from tomorrow onward. Then, the sender is unwilling to invest today in his reputation. For any such β,

$$\overline{V}(\beta) = \mu_0[(1-\delta) + \delta\overline{V}(\beta)] + (1-\mu_0)[(1-\delta) + \delta\overline{V}(\beta_{H,\ell})] > \frac{\delta - \mu_0(2\delta - 1)}{1 - \delta\mu_0}.$$

Recall the definition of $\bar{\beta}$. It is easy to verify that for any $\beta \geq \bar{\beta}$, the sender always says H.

Proof of Lemma 5: First, we show that for any $\beta \in [\bar{\beta}, 1)$, $\pi(\beta) = 1$ and $\alpha(\beta) = 1$. Assume by contradiction that for some $\beta \geq \bar{\beta}$, $\pi(\beta) < 1$ so that the sender recommends L with positive probability when $\theta = \ell$. Then $v_\ell \leq \delta\overline{V}(\beta_{L,\ell}) < \delta$ (because $\overline{V}$ is strictly increasing and $\overline{V}(1) = 1$) and $v_h \leq \alpha(\beta)(1-\delta) + \delta\overline{V}(\beta)$. Since $\alpha(\beta) \leq 1$,

$$\overline{V}(\beta) < \mu_0[(1-\delta) + \delta\overline{V}(\beta)] + (1-\mu_0)\delta$$

$$\iff \quad \overline{V}(\beta) < \frac{\delta - \mu_0(2\delta - 1)}{1 - \delta\mu_0},$$

which is a contradiction. Therefore, $\pi(\beta) = 1$ for all $\beta \geq \bar{\beta}$. Note that when $\pi(\beta) = 1$, $\beta_{H,\ell} = \beta_{H,\ell}(\beta, 1) < \beta$. If $\alpha(\beta) = 0$ for some $\beta \geq \bar{\beta}$, then

$$\overline{V}(\beta) \leq \mu_0\delta\overline{V}(\beta) + (1-\mu_0)\delta\overline{V}(\beta_{H,\ell}) < \delta\overline{V}(\beta),$$

[20]In this sense, $\alpha(\beta) = 1$ is "dominant" for the receiver in this range of β's, suggesting our notation β_d.

which is a contradiction. Therefore, $\alpha(\beta) > 0$ and it must be optimal for the receiver to follow the recommendation H even though she knows that $\pi(\beta) = 1$. That is, $\mu(1, \beta) \geq 1/2$ for all $\beta \geq \bar{\beta}$, and thus $\bar{\beta} \geq \beta_d$. But, for $\beta > \beta_d$, $\mu(1, \beta) > 1/2$ and the receiver strictly prefers to accept the recommendation H. That is, $\alpha(\beta) = 1$ and we can set $w_{H,h} = \overline{V}(\beta)$ and $w_{H,\ell} = \overline{V}(\beta_{H,\ell})$ to get

$$\overline{V}(\beta) = \mu_0[(1-\delta) + \delta\overline{V}(\beta)] + (1-\mu_0)[(1-\delta) + \delta\overline{V}(\beta_{H,\ell})].$$

Since $\overline{V}$ is continuous, taking the limit of the above equality as $\beta \downarrow \bar{\beta}$, gives

$$\overline{V}(\bar{\beta}_{H,\ell}) = 2 - \frac{1}{\delta} \quad \text{where} \quad \bar{\beta}_{H,\ell} = \beta_{H,\ell}(\bar{\beta}, 1).$$

Given that $\overline{V}(\bar{\beta}) = [\delta - \mu_0(2\delta - 1)]/[1 - \delta\mu_0]$ and $\overline{V}(\bar{\beta}_{H,\ell}) = 2 - 1/\delta$, we must also have that $\alpha(\bar{\beta}) = 1$. Thus $\alpha(\beta) = 1$ for all $\beta \in [\bar{\beta}, 1]$.

Second, we prove that for any $\beta < \bar{\beta}$, $\pi(\beta) < 1$. By contradiction, assume that $\beta < \bar{\beta}$ and $\pi(\beta) = 1$. Then, $\beta_{H,\ell} = \beta_{H,\ell}(\beta, 1) < \bar{\beta}_{H,\ell} = \beta_{H,\ell}(\bar{\beta}, 1)$ and $v_\ell \leq \alpha(\beta)(1-\delta) + \delta\overline{V}(\beta_{H,\ell}) < 1 - \delta + \delta\overline{V}(\bar{\beta}_{H,\ell}) = \delta$. But if out of equilibrium the sender recommends L after $\theta = \ell$ (an action that only the behavioral type takes after $\theta = \ell$), then $\beta_{L,\ell} = 1$ and his continuation value is $\delta\overline{V}(1) = \delta$, so recommending H for sure is not incentive compatible for the sender. □

Proof of Lemma 6: Let $\beta \in [0, \overline{\beta})$. By Lemmas 3, 4 and 5 and by definition of admissibility from Section 3.2, the tuple $(\pi(\beta), \alpha(\beta), w)$ that supports $\overline{V}(\beta)$ must be an optimal solution of

$$\begin{aligned}
\overline{V}(\beta) \;=\; \max_{(\pi,\alpha,w)} \quad & \mu_0[\alpha(1-\delta) + \delta w_{H,h}] + (1-\mu_0)\delta w_{L,\ell} \\
\text{s.t.} \quad & \pi_B \leq \pi \leq \min\{\bar{\pi}(\beta), 1\} \text{ and } \alpha \in [0,1], \\
& w_{H,h} \in V(\beta) \text{ and } w_{m,\ell} \in V(\beta_{m,\ell}(\beta, \pi)) \text{ for } m = H, L, \\
& \delta[w_{L,\ell} - w_{H,\ell}] = \alpha(1-\delta), \qquad \text{(IC)} \\
& \alpha \text{ is a best reply to } \pi.
\end{aligned}$$

Substituting $\alpha = \delta[w_{L,\ell} - w_{H,\ell}]/(1-\delta)$ from (IC) into the objective function, we get

$$\begin{aligned}
\overline{V}(\beta) \;=\; \max_{(\pi,w)} \quad & \mu_0\delta\left[\frac{1}{\mu_0}w_{L,\ell} - w_{H,\ell} + w_{H,h}\right] \\
\text{s.t.} \quad & \pi_B \leq \pi \leq \min\{\bar{\pi}(\beta), 1\}, \\
& w_{H,h} \in V(\beta) \text{ and } w_{m,\ell} \in V(\beta_{m,\ell}(\beta, \pi)) \text{ for } m = H, L, \\
& \delta[w_{L,\ell} - w_{H,\ell}] \leq (1-\delta) \qquad \text{(BC)} \\
& [1 - \delta(w_{L,\ell} - w_{H,\ell})/(1-\delta)][\bar{\pi}(\beta) - \pi] = 0 \qquad \text{(CC)}
\end{aligned}$$

The complementarity condition (CC) ensures that $\alpha = \delta(w_{L,\ell} - w_{H,\ell})/(1-\delta)$ is a best reply to π: if $\pi < \bar{\pi}(\beta)$, the receiver strictly prefers to accept the H recommendation, so α must be equal to 1. The (BC) constraint ensures that $\alpha \leq 1$.

Increasing the value of $w_{H,h}$ increases the value of $\overline{V}(\beta)$ and strengthens the incentives for the sender to recommend H when $\theta = h$. Therefore $w_{H,h} = \overline{V}(\beta)$. Now, if $w_{H,\ell} < \overline{V}(\beta_{H,\ell})$ and $w_{L,\ell} < \overline{V}(\beta_{L,\ell})$, then these two continuation values could be increased by the same small amount. This would increase the value of $\overline{V}(\beta)$ while keeping incentives at $\theta = \ell$ the same. Therefore, it must be that $w_{H,\ell} = \overline{V}(\beta_{H,\ell})$ or $w_{L,\ell} = \overline{V}(\beta_{L,\ell})$.

Assume by contradiction that $w_{L,\ell} < \overline{V}(\beta_{L,\ell})$ and $w_{H,\ell} = \overline{V}(\beta_{H,\ell})$. If the (IC) constraint were slack (so $\alpha(\beta) < 1$), we could increase $w_{L,\ell}$ to improve $\overline{V}(\beta)$. Therefore, it must be that the (BC) constraint is tight (i.e., $\alpha(\beta) = 1$ and $w_{L,\ell} = w_{H,\ell} + (1-\delta)/\delta$). But consider adjusting $(\pi(\beta), w)$ as follows. Decrease $\pi(\beta)$ a little to obtain $\tilde{\pi}(\beta)$. This change will produce new posteriors $\tilde{\beta}_{H,\ell}$ and $\tilde{\beta}_{L,\ell}$, where $\beta_{H,\ell} < \tilde{\beta}_{H,\ell} < \tilde{\beta}_{L,\ell} < \beta_{L,\ell}$. Let $\tilde{w}_{H,h} = w_{H,h}$, $\tilde{w}_{H,\ell} = \overline{V}(\tilde{\beta}_{H,\ell})$ and $\tilde{w}_{L,\ell} = \tilde{w}_{H,\ell} + (1-\delta)/\delta$. The new $(\tilde{\pi}(\beta), \tilde{w})$ is also feasible and delivers a higher value than $(\pi(\beta), w)$. This is a contradiction.

Finally we show that if $\pi(\beta) < \bar{\pi}(\beta)$, it must be that $\alpha(\beta) = 1$ and $w_{H,\ell} = \overline{V}(\beta_{H,\ell})$. If $\pi(\beta) < \bar{\pi}(\beta)$, (CC) implies that $\alpha(\beta) = 1$ and (BC) is tight. But if $w_{H,\ell} < \overline{V}(\beta_{H,\ell})$, we can increase $\pi(\beta)$ a bit to $\tilde{\pi}(\beta)$ and adjust continuation values as follows. Make $\tilde{w}_{H,h} = w_{H,h}$, $\tilde{w}_{L,\ell} = \overline{V}(\tilde{\beta}_{L,\ell})$, and $\tilde{w}_{H,\ell} = \tilde{w}_{L,\ell} - (1-\delta)/\delta$. Since $\tilde{\beta}_{L,\ell} > \beta_{L,\ell}$, $\tilde{w}_{L,\ell} > w_{L,\ell}$ and $(\tilde{\pi}(\beta), \tilde{w})$ delivers a higher value. A contradiction. □

Before we state and prove Lemmas 7 and 8, it is convenient to introduce some definitions. When $\pi = \bar{\pi}(\beta)$ the posteriors $\beta_{H,\ell}(\beta, \pi)$ and $\beta_{L,\ell}(\beta, \pi)$ take a particularly simple form:

$$\beta_{H,\ell}(\beta, \bar{\pi}(\beta)) = \lambda_0 \beta, \quad \beta_{L,\ell}(\beta, \bar{\pi}(\beta)) = \lambda_1 \beta,$$

where

$$\lambda_0 = \frac{\pi_B}{\pi^*} = \frac{\pi_B(1-\mu_0)}{\mu_0} \quad \text{and} \quad \lambda_1 = \frac{1-\pi_B}{1-\pi^*} = \frac{(1-\pi_B)(1-\mu_0)}{1-2\mu_0}.$$

Since $\pi_B < \pi^* < 1/2$, $0 < \lambda_0 < 1 < \lambda_1$ and $\lambda_0 \lambda_1 < 1$.

Lemma 7. *For $\beta \in [0, \overline{\beta})$, $\beta_{H,\ell}(\beta, \pi(\beta))$ and $\beta_{L,\ell}(\beta, \pi(\beta))$ are increasing functions of β.*

Proof: Let $\beta \in [0, \overline{\beta})$ be such that for some $\epsilon > 0$, $\pi(\beta') = \bar{\pi}(\beta')$ for all $\beta' \in [\beta, \beta + \epsilon]$. Then, for any $\beta' \in (\beta, \beta + \epsilon]$

$$\beta_{H,\ell}(\beta, \bar{\pi}(\beta)) = \lambda_0 \beta < \lambda_0 \beta' = \beta_{H,\ell}(\beta', \bar{\pi}(\beta')) \quad \text{and}$$
$$\beta_{L,\ell}(\beta, \bar{\pi}(\beta)) = \lambda_1 \beta < \lambda_1 \beta' = \beta_{L,\ell}(\beta', \bar{\pi}(\beta')).$$

Now consider $\beta \in [0, \overline{\beta})$ such that for some $\epsilon > 0$, $\pi(\beta') < \bar{\pi}(\beta')$ for all $\beta' \in (\beta, \beta + \epsilon)$. Assume by contradiction that for some $\beta' \in (\beta, \beta + \epsilon)$ we have that $\beta_{H,\ell}(\beta, \pi(\beta)) \geq \beta_{H,\ell}(\beta', \pi(\beta'))$. This implies that $\pi(\beta') > \pi(\beta)$. But $\beta' > \beta$ and $\pi(\beta') > \pi(\beta)$ imply that $\beta_{L,\ell}(\beta, \pi(\beta)) < \beta_{L,\ell}(\beta', \pi(\beta'))$. Since $\overline{V}$ is a strictly increasing function,

$$\begin{aligned} &\delta[\overline{V}(\beta_{L,\ell}(\beta', \pi(\beta'))) - \overline{V}(\beta_{H,\ell}(\beta', \pi(\beta')))] \\ &> \delta[\overline{V}(\beta_{L,\ell}(\beta, \pi(\beta))) - \overline{V}(\beta_{H,\ell}(\beta, \pi(\beta)))] = 1 - \delta, \end{aligned}$$

a contradiction. Therefore $\beta_{H,\ell}(\beta, \pi(\beta)) < \beta_{H,\ell}(\beta', \pi(\beta'))$ for all $\beta' \in (\beta, \beta + \epsilon)$. Since

$$\delta[\overline{V}(\beta_{L,\ell}(\beta', \pi(\beta'))) - \overline{V}(\beta_{H,\ell}(\beta', \pi(\beta')))] = 1 - \delta$$

and $\overline{V}$ is strictly increasing, it must be that $\overline{V}(\beta_{L,\ell}(\beta', \pi(\beta'))) > \overline{V}(\beta_{L,\ell}(\beta, \pi(\beta)))$ and $\beta_{L,\ell}(\beta', \pi(\beta')) > \beta_{L,\ell}(\beta, \pi(\beta))$. □

For each reputation β, let $\pi(\beta)$ and $\alpha(\beta)$ be the strategies of the sender and the receiver in the first period of the sender-preferred equilibrium starting at β. Fixing β_0, let $\pi_0 = \pi(\beta_0)$,

$$\beta_H = \beta_{H,\ell}(\beta_0, \pi_0) \quad \text{and} \quad \beta_L = \beta_{L,\ell}(\beta_0, \pi_0).$$

Section 5.2 considers the reputation maintenance strategy based at β_0. By Lemma 7, $\beta_H \leq \beta_{H,\ell}(\beta, \pi(\beta)) < \beta$ for all $\beta \in [\beta_0, \beta_L]$ and $\beta < \beta_{L,\ell}(\beta, \pi(\beta)) < \beta_L$ for all $\beta \in [\beta_H, \beta_0)$. Also $\beta_{H,h}(\beta, \pi(\beta)) = \beta$. Hence, when the sender follows this strategy, his reputation stays in the interval $[\beta_H, \beta_L]$ in every period.

In the spirit of dynamic programming, for an arbitrary continuation value function W from $[\beta_H, \beta_L]$ to $[0, 1]$, compute for each $\beta \in [\beta_H, \beta_L]$ the value of doing one round of reputation maintenance (that is, play L if $\beta \leq \beta_0$ and $\theta = \ell$, play H otherwise), fixing $\alpha(\beta) = 1$ and using continuation values given by W.

Formally, let $\mathbb{W} = [0, 1]^{[\beta_H, \beta_L]}$ be the set of all functions W from $[\beta_H, \beta_L]$ to $[0, 1]$. Endow $\mathbb{W}$ with the sup norm:

$$||W|| = \sup \{|W(\beta)| \mid \beta \in [\beta_H, \beta_L]\}.$$

Let $T^\pi : \mathbb{W} \to \mathbb{W}$ be the map

$$T^\pi(W)(\beta) = \begin{cases} \mu_0(1 - \delta) + \delta[\mu_0 W(\beta) + (1 - \mu_0) W(\beta_{L,\ell}(\beta, \pi(\beta)))] & \text{for } \beta \in [\beta_H, \beta_0) \\ (1 - \delta) + \delta[\mu_0 W(\beta) + (1 - \mu_0) W(\beta_{H,\ell}(\beta, \pi(\beta)))] & \text{for } \beta \in [\beta_0, \beta_L]. \end{cases}$$

One can easily check that T^π is a contraction and therefore it has a unique fixed point. Denote by $V^\pi(\beta, \delta)$ this fixed point to make explicit that it depends on δ as well.

Lemma 8. *$V^\pi(\beta,\delta) \geq \overline{V}(\beta,\delta)$ for all $\beta \in [\beta_H, \beta_L]$. Conversely, if $[\beta_0, \beta_L]$ is in Region 2 and $\alpha(\beta) = 1$ for all $\beta \in [\beta_H, \beta_0)$, then $\overline{V}(\beta,\delta) = V^\pi(\beta,\delta)$ for all $\beta \in [\beta_H, \beta_L]$.*

Proof: We first show that if $W(\beta) \geq \overline{V}(\beta,\delta)$ for all $\beta \in [\beta_H, \beta_L]$, then $T^\pi(W)(\beta) \geq \overline{V}(\beta,\delta)$ for all $\beta \in [\beta_H, \beta_L]$. If $\beta \in [\beta_H, \beta_0)$, by right promise keeping we have that

$$\begin{aligned}\overline{V}(\beta,\delta) &= \alpha(\beta)\mu_0(1-\delta) + \delta\left[\mu_0\overline{V}(\beta,\delta) + (1-\mu_0)\overline{V}(\beta_{L,\ell},\delta)\right] \\ &\leq \mu_0(1-\delta) + \delta\left[\mu_0 W(\beta) + (1-\mu_0)W(\beta_{L,\ell})\right] = T^\pi(W)(\beta).\end{aligned}$$

If $\beta \in [\beta_0, \beta_L]$, by left promise keeping there exists $w_{H,\ell} \leq \overline{V}(\beta_{H,\ell},\delta)$ such that

$$\begin{aligned}\overline{V}(\beta,\delta) &= \alpha(\beta)(1-\delta) + \delta\left[\mu_0\overline{V}(\beta,\delta) + (1-\mu_0)w_{H,\ell}\right] \\ &\leq (1-\delta) + \delta\left[\mu_0 W(\beta) + (1-\mu_0)W(\beta_{H,\ell})\right] = T^\pi(W)(\beta).\end{aligned}$$

Since $\{W \in \mathbb{W} \mid W(\beta) \geq \overline{V}(\beta,\delta) \text{ for all } \beta \in [\beta_H, \beta_L]\}$ is a closed set in $(\mathbb{W}, ||\cdot||)$, $V^\pi(\beta,\delta) \geq \overline{V}(\beta,\delta)$ for all $\beta \in [\beta_H, \beta_L]$.

Now, assume that $\alpha(\beta) = 1$ for all $\beta \in [\beta_H, \beta_0)$ and that $[\beta_0, \beta_L]$ is in Region 2. The former implies that for all $\beta \in [\beta_H, \beta_0)$, $\overline{V}(\beta,\delta) = T^\pi(\overline{V}(\cdot,\delta))(\beta)$. The latter implies that for all $\beta \in [\beta_0, \beta_L]$, $\alpha(\beta) = 1$ and $w_{H,\ell} = \overline{V}(\beta_{H,\ell},\delta)$, so that $\overline{V}(\beta,\delta) = T^\pi(\overline{V}(\cdot,\delta))(\beta)$. That is $\overline{V}(\cdot,\delta)$ is a fixed point of T^π, and therefore $\overline{V} = V^\pi$. □

Remark: If $\beta_H > \beta_d$ then $\bar{\pi}(\beta) > 1$ for all $\beta \geq \beta_H$. Therefore $\pi(\beta) < \bar{\pi}(\beta)$ and $\alpha(\beta) = 1$ for all $\beta \in [\beta_H, \beta_L]$.

Similarly, one can easily show that if W is continuous, $T^\pi(W)$ is continuous, and if W is weakly increasing so is $T^\pi(W)$. Hence, $V^\pi(\beta,\delta)$ is continuous and weakly increasing in β.

Lemma 9 below establishes that when the sender follows the reputation maintenance strategy, the ratio of the frequencies with which he recommends H and with which he recommends L in periods when $\theta = \ell$ is roughly

$$R(\pi_0) = \log\left[\frac{1-\pi_B}{1-\pi_0}\right] \Big/ \log\left[\frac{\pi_0}{\pi_B}\right].$$

and therefore when δ is close to 1 (so the order is which the recommendations of H and L happen does not affect the sender's payoff too much), $V^\pi(\beta_0,\delta)$ is approximately equal to

$$V^M(\pi_0) = \mu_0 + (1-\mu_0)\frac{R(\pi_0)}{R(\pi_0)+1}.$$

Lemma 9 makes the counterfactual assumption that $\pi(\beta)$ remains constant in $[\beta_H, \beta_L]$. This eliminates one dimension of variation which allows us to provide a much simpler analysis and proof that abstracts from the complexities involved in the proof of Lemma 10.

Lemma 9. *(Stationary Promise-keeping) Fix β_0 and let β_H and β_L be as defined above. Assume $\pi(\beta) = \pi_0$ for all $\beta \in [\beta_H, \beta_L]$ and that the sender follows the reputation maintenance strategy. Suppose that $[\beta_0, \beta_L]$ is contained in Region. Then there exists a constant $D > 0$ (independent of δ and π_0) such that*

$$|\overline{V}(\beta_0, \delta) - V^M(\pi_0)| \leq D \cdot (1 - \delta).$$

Proof: This is a Corollary of Lemma 10. We provide a sketch of the proof here, because it is a simpler introduction to how stationary promise-keeping works.

Suppose that for the first $n+m$ instances when $\theta = \ell$, the sender has recommended H in n periods and L in m periods. We first show that in this case

$$\frac{n-1}{m} \leq R(\pi_0) \leq \frac{n}{m-1}.$$

By Baye's rule, the receiver's posterior after the first $n+m$ instances of $\theta = \ell$ is

$$\hat{\beta} = \frac{\beta \pi_B^n (1-\pi_B)^m}{\beta \pi_B^n (1-\pi_B)^m + (1-\beta)\pi_0^n (1-\pi_0)^m} = \frac{\beta}{\beta + (1-\beta) L_{n,m}} \qquad (*)$$

$$\text{where} \quad L_{n,m} = \frac{\pi_0^n (1-\pi_0)^m}{\pi_B^n (1-\pi_B)^m} \quad \text{is the likelihood ratio.}$$

Since $\hat{\beta} \in [\beta_H, \beta_L]$,

$$\frac{\pi_0}{\pi_B} = L_{1,0} \geq L_{n,m} \geq L_{0,1} = \frac{1-\pi_0}{1-\pi_B} \quad \Longleftrightarrow \quad \frac{n-1}{m} \leq R \leq \frac{n}{m-1}$$
$$\Longleftrightarrow \quad \frac{n-1}{n+m-1} \leq \frac{R}{R+1} \leq \frac{n}{n+m-1}. \qquad (**)$$

When the sender follows the reputation maintenance strategy for a large number of periods k, roughly in about $k\mu_0$ periods the state is $\theta = h$ and the sender sends message H; in about $k(1-\mu_0)\frac{R}{R+1}$ periods the state is $\theta = \ell$ and the sender sends the message H, and in about $k(1-\mu_0)\frac{1}{R+1}$ periods the state is $\theta = \ell$ and the sender sends the message L. Since $\alpha(\beta) = 1$ for $\beta \in [\beta_H, \beta_L]$, the receiver accepts every H recommendation. When δ is close to 1, it does not matter much in which order these events occur and the sender collects in the first k periods a total discounted payoff approximately equal to

$$(1-\delta)\left[\mu_0 + (1-\mu_0)\frac{R}{R+1}\right](1 + \delta + \cdots + \delta^{k-1}) = \left[\mu_0 + (1-\mu_0)\frac{R}{R+1}\right](1-\delta^k).$$

Taking the limit as $k \to \infty$, we obtain that $\lim_{\delta \to 1} \overline{V}(\beta_0, \delta) = V^M(\pi_0)$. □

Lemma 10. *Fix δ close to 1 and β_0. Let $\pi_0 = \pi(\beta_0)$ and (β_H, β_L) be as defined above. Assume that for some $\epsilon > 0$, $\pi(\beta_H) \geq \pi_0 - \epsilon$ and $\pi(\beta_L) \leq \pi_0 + \epsilon$. Then*

$$V^M(\pi_0) + O([1-\delta]^2) - E(\pi_0)\epsilon \leq V^\pi(\beta_0, \delta) \leq V^M(\pi_0) + O([1-\delta]^2),$$

where $E(\pi_0)$ is a continuous and strictly positive function from $[\pi^*, 1)$ to $\mathbb{R}$, and the function O is such that $\lim_{x\to 0} O(x)/x < \infty$.

Proof: Assume that δ is close to 1 and that the sender follows the reputation maintenance strategy starting at β_0. Since $\pi(\beta)$ is not constant and equal to $\pi(\beta_0) = \pi_0$ for all $\beta \in [\beta_H, \beta_L]$, $\overline{V}(\beta_0, \delta)$ will typically differ from $V^M(\pi_0)$. Let $\{\beta_k\}$ be the (random) sequence of posteriors generated along the way. Then

$$\beta_{k+1} = \begin{cases} \beta_k & \text{when } \theta_k = h \\ \beta_{L,\ell}(\beta_k, \pi(\beta_k)) & \text{when } \theta_k = \ell \text{ and } \beta_k < \beta_0 \\ \beta_{H,\ell}(\beta_k, \pi(\beta_k)) & \text{when } \theta_k = \ell \text{ and } \beta_k \geq \beta_0. \end{cases}$$

This implies that

$$\beta_k = \frac{\beta_0}{\beta_0 + (1-\beta_0)L_k} \quad \text{where}$$
$$L_k = \left[\prod_{\{j<k|\theta_j = \ell \text{ and } \beta_j \geq \beta_0\}} \frac{\pi(\beta_j)}{\pi_B}\right] \times \left[\prod_{\{j<k|\theta_j = \ell \text{ and } \beta_j < \beta_0\}} \frac{1-\pi(\beta_j)}{1-\pi_B}\right].$$

For any k, let $n = |\{j < k \mid \theta_j = \ell \text{ and } \beta_j \geq \beta_0\}|$ and $m = |\{j < k \mid \theta_j = \ell \text{ and } \beta_j < \beta_0\}|$. Since $\pi_0 \geq \pi(\beta_j) \geq \pi(\beta_H) \geq \pi_0 - \epsilon$ for any $\beta_j \in [\beta_H, \beta_0)$, and $\pi_0 \leq \pi(\beta_j) \leq \pi(\beta_L) \leq \pi_0 + \epsilon$ for any $\beta_j \in [\beta_0, \beta_L]$, we get that

$$\left[\frac{1-\pi_0}{1-\pi_B}\right]^m \left[\frac{\pi_0}{\pi_B}\right]^n \leq L_k \leq \left[\frac{1-\pi_0+\epsilon}{1-\pi_B}\right]^m \left[\frac{\pi_0+\epsilon}{\pi_B}\right]^n.$$

Since $\beta_k \in [\beta_H, \beta_L]$ for all k, it must be that $(1-\pi_0)/(1-\pi_B) \leq L_k \leq \pi_0/\pi_B$. Therefore,

$$\left[\frac{1-\pi_0}{1-\pi_B}\right]^m \left[\frac{\pi_0}{\pi_B}\right]^n \leq \frac{\pi_0}{\pi_B} \quad \text{and} \quad \frac{1-\pi_0}{1-\pi_B} \leq \left[\frac{1-\pi_0+\epsilon}{1-\pi_B}\right]^m \left[\frac{\pi_0+\epsilon}{\pi_B}\right]^n.$$

The first inequality implies that

$$\frac{n-1}{m} \leq R(\pi_0) \quad \text{or} \quad \frac{n-1}{n+m-1} \leq \frac{R(\pi_0)}{R(\pi_0)+1}, \tag{$*$}$$

and the second inequality implies that

$$n \log\left[\frac{\pi_0+\epsilon}{\pi_B}\right] \geq m \log\left[\frac{1-\pi_B}{1-\pi_0+\epsilon}\right] - \log\left[\frac{1-\pi_B}{1-\pi_0}\right].$$

By concavity and convexity in ϵ of the corresponding coefficients, we have that

$$\log\left[\frac{\pi_0+\epsilon}{\pi_B}\right] \leq C(\pi_0) + \frac{\epsilon}{\pi_0} \quad \text{and} \quad \log\left[\frac{1-\pi_B}{1-\pi_0+\epsilon}\right] \geq D(\pi_0) - \frac{\epsilon}{1-\pi_0}, \quad \text{where}$$
$$C(\pi_0) = \log\left[\frac{\pi_0}{\pi_B}\right] \quad \text{and} \quad D(\pi_0) = \log\left[\frac{1-\pi_B}{1-\pi_0}\right].$$

Note that $D(\pi_0)/C(\pi_0) = R(\pi_0)$, so the previous inequality implies that

$$\frac{n}{m-1} \geq \frac{D(\pi_0) - \frac{m}{m-1}\frac{\epsilon}{1-\pi_0}}{C(\pi_0) + \frac{\epsilon}{\pi_0}} \geq R(\pi_0) - E(\pi_0)\epsilon \quad \text{where}$$
$$E(\pi_0) = \left[\frac{2C(\pi_0)}{1-\pi_0} + \frac{D(\pi_0)}{\pi_0}\right] \Big/ C(\pi_0)^2.$$

Therefore

$$\frac{n}{n+m-1} \geq \frac{R(\pi_0)}{R(\pi_0)+1} - E(\pi_0)\epsilon. \tag{$**$}$$

When the sender follows the reputation maintenance strategy, the sequence of recommendations in the set of periods j where $\theta_j = \ell$ is deterministic. To compute the value of this strategy, we consider the following accounting system. For fixed n (large), stop when the sender recommends H for the n-th time in a period in which the state is ℓ. This will include n periods such that $(\theta_j, a_j) = (\ell, H)$ (including the last), a deterministic number m_1 of periods such that $(\theta_j, a_j) = (\ell, L)$, and a random number k of periods such that $(\theta_j, a_j) = (h, H)$. Let V_1 be the expected discounted value of the payoffs the sender accumulates until the process is stopped. Let E_k be the event such that at the time the process stops, there have been exactly k periods in which the state is h. Though m_1 is a deterministic function of n, a precise expression for m_1 is hard to obtain. However, the previous analysis places strict bounds on m_1 given by $(*)$ and $(**)$.

$$V_1 = \sum_{k=0}^{\infty} \binom{n+m_1-1+k}{k} \mu_0^k (1-\mu_0)^{n+m_1} \mathbb{E}[(1-\delta)S_k]$$
$$\text{where} \quad S_k = \sum_{j=0}^{n+m_1-1+k} \delta^j \mathbb{1}_{\{a_j=H\}}$$

Though the periods where $a_j = H$ are random, we can bound $\mathbb{E}[(1-\delta)S_k]$ easily, assuming in two extreme cases that they all occur at the beginning or that they all occur at the end:

$$(1-\delta)(\delta^{m_1} + \delta^{m_1+1} + \cdots + \delta^{m_1+k+n-1}) = \delta^{m_1}(1-\delta^{k+n})$$
$$\leq \mathbb{E}[(1-\delta)S_k] \leq (1-\delta)(1+\delta+\cdots+\delta^{k+n-1}) = (1-\delta^{k+n}).$$

Note that

$$\sum_{k=0}^{\infty} \binom{n+m_1-1+k}{k} \mu_0^k (1-\mu_0)^{n+m_1} \delta^{k+n}$$
$$= (1-\mu_0)^{n+m_1}\delta^n \sum_{k=0}^{\infty} \binom{n+m_1-1+k}{k} (\delta\mu_0)^k = \frac{1}{\delta^{m_1}} \left(\frac{(1-\mu_0)\delta}{1-\delta\mu_0}\right)^{n+m_1}.$$

Let $\Delta_1 = \left(\frac{(1-\mu_0)\delta}{1-\delta\mu_0}\right)^{n+m_1}$. Replacing these bounds in the computation of V_1 we obtain:

$$\delta^{m_1} - \Delta_1 \leq V_1 \leq 1 - \frac{\Delta_1}{\delta^{m_1}}.$$

Finally note that

$$\mathbb{E}[\delta^{n+m_1+k}] = \Delta_1.$$

Having computed (bounds for) V_1, let us restart the process and stop it again when for the second time the sender accumulates n periods where $(\theta_j, a_j) = (\ell, H)$. Again, this will include a deterministic number of periods m_2 where $(\theta_j, a_j) = (\ell, L)$ and a random number periods k where $\theta_j = h$. Let V_2 be the expected discounted value of the payoffs that the sender accumulates between the first and the second time the process is stopped. Define similarly m_j and V_j for $j \geq 3$. Then

$$V^\pi(\beta_0, \delta) = V_1 + \Delta_1[V_2 + \Delta_2[V_3 + \Delta_3[\ldots]]].$$

For any $m \in \mathbb{R}_+$, let $\Delta(m,\delta) = \left(\frac{(1-\mu_0)\delta}{1-\delta\mu_0}\right)^{n+m}$. By continuity, there exists $m \in [\underline{m}, \overline{m}]$, where $\underline{m} = \min\{m_j\}$ and $\overline{m} = \max\{m_j\}$, such that

$$V^\pi(\beta_0,\delta) \geq (\delta^{m_1} - \Delta(m_1,\delta)) + \Delta(m_1,\delta)[(\delta^{m_2} - \Delta(m_2,\delta)) + \Delta(m_2,\delta)[\cdots]]$$
$$= (\delta^m - \Delta(m,\delta))(1 + \Delta(m,\delta) + \Delta(m,\delta)^2 + \cdots) = \frac{\delta^m - \Delta(m,\delta)}{1-\Delta(m,\delta)}.$$

By Taylor series expansion

$$\Delta(m,\delta) = \Delta(m,1) + \Delta_\delta(m,1)(\delta-1) + O((1-\delta)^2)$$
$$= 1 + \frac{1}{1-\mu_0}(n+m)(\delta-1) + O([1-\delta]^2) \quad \text{and}$$
$$\delta^m = 1 + m(\delta-1) + O([1-\delta]^2).$$

Therefore, inequality (∗∗) implies that

$$\begin{aligned} V^\pi(\beta_0,\delta) \geq & \ \frac{-m+\frac{1}{1-\mu_0}(n+m)}{\frac{1}{1-\mu_0}(n+m)} + O([1-\delta]^2) \\ = & \ \mu_0 + (1-\mu_0)\frac{n}{n+m} + O([1-\delta]^2) \\ \geq & \ V^M(\pi_0) - \frac{n(1-\mu_0)}{(n+m)(n+m-1)} - E(\pi_0)\epsilon + O([1-\delta]^2). \end{aligned}$$

Since n is arbitrary, we can make the second term on the right hand side arbitrarily small by choosing n large enough. This establishes the lower bound. The upper bound is established similarly. Here we note that by Taylor series expansion,

$$\Delta(m,\delta)/\delta^m = 1 + [-m + \Delta_\delta(m,1)](\delta-1) + O([1-\delta]^2).$$

Then, inequality (∗) implies that

$$\begin{aligned} V^\pi(\beta_0,\delta) &\leq \mu_0 + (1-\mu_0)\frac{n}{n+m} + O([1-\delta]^2) \\ &\leq V^M(\pi_0) + \frac{1-\mu_0}{n+m-1} + O([1-\delta]^2), \end{aligned}$$

and again the second term is made arbitrarily small by choosing n large enough. □

By definition of $O([1-\delta]^2)$, there exists $\underline{\delta} < 1$ such that, under the assumptions of Lemma 10,

$$V^\pi(\beta_0,\delta) \geq V^M(\pi_0) - (1-\delta) - E(\pi_0)\epsilon \quad \text{for all } \delta \in [\underline{\delta}, 1).$$

Corollary 2. *Fix β_0 and $\delta \in [\underline{\delta}, 1)$. Let $\pi_0 = \pi(\beta_0)$ and (β_H, β_L) be as defined above. If*

$$V^\pi(\beta_0,\delta) < V^M(\pi_0) - (1-\delta) - E(\pi_0)\epsilon$$

for some $\epsilon > 0$, then $\epsilon < \max\{\pi_0 - \pi(\beta_H), \pi(\beta_L) - \pi_0\}$.

Proof of Lemma 11: Fix $\epsilon \in (0, 1/2)$ and pick any $\delta > \delta_1 \equiv [2(1-\mu_0) - \epsilon]/[2(1-\mu_0) - \epsilon\mu_0]$. Then

$$\overline{V}(\overline{\beta},\delta) - (1-\epsilon) = \frac{\delta - \mu_0(2\delta-1)}{1-\delta\mu_0} - (1-\epsilon) > \epsilon/2.$$

If $\overline{V}(\epsilon,\delta) > 1-\epsilon$, then $\overline{V}(1-\epsilon,\delta) - \overline{V}(\epsilon,\delta) < 1-(1-\epsilon) = \epsilon$ and we are done. Hereafter, assume that $\overline{V}(\epsilon,\delta) \leq 1-\epsilon$.

Set $\beta_0 = \epsilon$ and inductively define $v_k = \overline{V}(\beta_k, \delta)$ and $\beta_{k+1} = \beta_{L,\ell}(\beta_k, \pi(\beta_k))$, $k \geq 0$. Then

$$\beta_k = \frac{\beta_0}{\beta_0 + (1-\beta_0)L_k} \quad \text{where} \quad L_k = \frac{1}{(1-\pi_B)^k}[(1-\pi(\beta_0))\cdots(1-\pi(\beta_{k-1}))].$$

Since $\pi(\beta) > \pi^* > \pi_B$ for all $\beta \in (0,1]$, $(1-\pi(\beta))/(1-\pi_B) < (1-\pi^*)/(1-\pi_B) < 1$ and

$$L_k < \left[\frac{1-\pi^*}{1-\pi_B}\right]^k.$$

Let K be the smallest integer such that $L_K \leq [\epsilon/(1-\epsilon)]^2$. Then $\beta_K > 1-\epsilon$. Since by assumption $v_0 = \overline{V}(\epsilon, \delta) \leq 1-\epsilon < \frac{\delta-\mu_0(2\delta-1)}{1-\delta\mu_0} = \overline{V}(\overline{\beta}, \delta)$, we have that $\beta_0 < \overline{\beta}$. For as long as $\beta_{k-1} < \bar{\beta}$, by right promise keeping, we have that

$$v_k < v_{k-1} + \frac{1-\delta}{\delta} \quad \text{so} \quad v_k < v_0 + k\frac{1-\delta}{\delta}.$$

Let δ_2 be such that $K(1-\delta_2)/\delta_2 = \epsilon/2$. Then, for any $\delta \geq \hat{\delta} = \max\{\delta_1, \delta_2\}$, $v_K - v_0 \leq \epsilon/2$. This implies that $v_k < \overline{V}(\overline{\beta}, \delta)$, and hence $\beta_k < \overline{\beta}$, for all $k \leq K$. Therefore $\overline{V}(1-\epsilon, \delta) - \overline{V}(\epsilon, \delta) < v_K - v_0 \leq \epsilon/2$. □

Define

$$\kappa = \frac{1-\mu_0\delta}{(1-\mu_0)\delta}, \quad \rho = \frac{\log(\kappa)}{\log(\lambda_1)}, \quad \ell_i = \log(\lambda_i) \quad i = 0,1,$$
$$V^{1,2}(\pi_B, \delta) = \mu_0 + \frac{1-\delta}{\delta[\kappa - \kappa^{\ell_0/\ell_1}]}.$$

Below, we will usually omit the variables in $V^{1,2}$; similarly we have omitted the variables in the definitions of κ and ρ. As stated in Lemma 13 below, $V^{1,2}(\pi_B, \delta)$ is approximately the value of $\overline{V}(\beta)$ when the posterior β is at the boundary between Region 1 and Region 2.

Lemma 12. *Assume Region 1b is empty. Then there exists a continuous function $a(\cdot)$ with the cyclical property that $a(\lambda_1\beta) = a(\beta)$ such that*

$$\overline{V}(\beta) = \mu_0 + a(\beta)\beta^\rho$$

for all $\beta < \underline{\beta}$.

Proof: Assume $\beta_k := \lambda_1^k\beta_0$ is in Region 1 for $k = 1, \ldots, K$. Denote $v_k = \overline{V}(\beta_k)$. The right promising keeping constraint then becomes

$$v_k = \mu_0[(1-\delta) + \delta v_k] + (1-\mu_0)\delta v_{k+1} \quad \text{or} \quad v_{k+1} = \kappa v_k - \frac{\mu_0(1-\delta)}{(1-\mu_0)\delta}.$$

The solution to this linear difference equation is

$$v_k = \mu_0 + \hat{a}\kappa^k$$

for some constant $\hat{a} > 0$ ($\hat{a} = v_0 - \mu_0$). Note that $k = \log[\beta_k/\beta_0]/\ell_1$. Therefore

$$\overline{V}(\beta_k) = \mu_0 + a\beta_k^\rho,$$

where $a > 0$ is constant (determined by the initial condition $\overline{V}(\beta_0)$).

If starting at a different posterior $\tilde{\beta}_0$ all the points $\tilde{\beta}_k = \lambda_1^k \tilde{\beta}_0$ for $k = 1, \dots, \tilde{K}$ are in Region 1, then there exists another constant $\tilde{a}$ such that $\overline{V}(\tilde{\beta}_k) = \mu_0 + \tilde{a}\tilde{\beta}_k^\rho$. Since $\overline{V}(\beta)$ is a continuous and increasing function of β,

$$\overline{V}(\beta) = \mu_0 + a(\beta)\beta^\rho \quad \text{for all } \beta \text{ in Region 1,}$$

where $a(\beta)$ is a continuous function of β such that $a(\lambda_1\beta) = a(\beta)$ for all β. $\square$

Lemma 13. *Assume Region 1b is empty and let $\hat{\beta}$ be such that $\overline{V}(\hat{\beta}) = V^{1,2}(\pi_B, \delta)$. Then*

(i) there is $\underline{\beta} \in [\hat{\beta}/\lambda_1, \lambda_1\hat{\beta}]$ such that $\overline{V}(\lambda_1\underline{\beta}) - \overline{V}(\lambda_0\underline{\beta}) = (1-\delta)/\delta$;

(ii) the interval $[0, \underline{\beta})$ is contained in Region 1 and $\underline{\beta}$ is the first reputation in Region 2: $\overline{V}(\lambda_1\beta) - \overline{V}(\lambda_0\beta) < (1-\delta)/\delta$ for all $\beta < \underline{\beta}$;

(iii) for any $\delta \geq 1/(1+\mu_0)$, $|\overline{V}(\underline{\beta}) - V^{1,2}| \leq 3(1-\delta)$.

Proof: For any fixed $\hat{a} > 0$, let $\hat{V}(\beta) = \mu_0 + \hat{a}\beta^\rho$ and $\Delta(\beta) = \hat{V}(\lambda_1\beta) - \hat{V}(\lambda_0\beta)$. Then, (1) $\Delta(0) = 0$; (2) $\Delta(\beta)$ is increasing in β; and (3) $\Delta(\beta) = (1-\delta)/\delta$ if and only if $\hat{V}(\beta) = V^{1,2}(\pi_B, \delta)$.

Let $\hat{\beta}$ be such that $\overline{V}(\hat{\beta}) = V^{1,2}$, and define $\hat{a} = a(\hat{\beta})$ (defined in Lemma 12). If $a(\lambda_0\hat{\beta}) = \hat{a}$, then $\overline{V}(\lambda_0\hat{\beta}) = \hat{V}(\lambda_0\hat{\beta})$ and $\overline{V}(\lambda_1\hat{\beta}) - \overline{V}(\lambda_0\hat{\beta}) = \hat{V}(\lambda_1\hat{\beta}) - \hat{V}(\lambda_0\hat{\beta}) = (1-\delta)/\delta$, since $a(\lambda_1\hat{\beta}) = \hat{a}$. But typically $a(\lambda_0\hat{\beta}) \neq \hat{a}$. Assume that $\overline{V}(\lambda_0\hat{\beta}) < \hat{V}(\lambda_0\hat{\beta})$. Then $\overline{V}(\lambda_1\hat{\beta}) - \overline{V}(\lambda_0\hat{\beta}) > \hat{V}(\lambda_1\hat{\beta}) - \hat{V}(\lambda_0\hat{\beta}) = (1-\delta)/\delta$ and $\hat{\beta}$ is already in the interior of Region 2. Since $a(\beta)$ is continuous and makes a full "cycle" in the interval $[\hat{\beta}/\lambda_1, \hat{\beta}]$, there exists $\tilde{\beta} \in (\hat{\beta}/\lambda_1, \hat{\beta})$ such that $\overline{V}(\lambda_0\tilde{\beta}) = \tilde{V}(\lambda_0\tilde{\beta})$. Since $\tilde{\beta} < \hat{\beta}$, $\overline{V}(\tilde{\beta}) < V^*$ and $\overline{V}(\lambda_1\tilde{\beta}) - \overline{V}(\lambda_0\tilde{\beta}) = \tilde{V}(\lambda_1\tilde{\beta}) - \tilde{V}(\lambda_0\tilde{\beta}) < (1-\delta)/\delta$, and $\tilde{\beta}$ is in Region 1. That is, the transition between Region 1 and Region 2 (when $\overline{V}(\lambda_1\beta) - \overline{V}(\lambda_0\beta) = (1-\delta)/\delta$) must occur at some $\beta \in [\hat{\beta}/\lambda_1, \hat{\beta}]$. Similarly, when $\overline{V}(\lambda_0\hat{\beta}) > \hat{V}(\lambda_0\hat{\beta})$, the transition must occur at some $\beta \in [\hat{\beta}, \lambda_1\hat{\beta}]$.

In summary, there exists $\underline{\beta} \in [\hat{\beta}/\lambda_1, \lambda_1\hat{\beta}]$, where the first transition between Region 1 and Region 2 occurs: $\overline{V}(\lambda_1\underline{\beta}) - \overline{V}(\lambda_0\underline{\beta}) = (1-\delta)/\delta$.

Let $\hat{a} = a(\underline{\beta})$. Since $\underline{\beta} \geq \hat{\beta}/\lambda_1$, $\overline{V}(\underline{\beta}) \geq \overline{V}(\hat{\beta}/\lambda_1)$. Since $\lambda_1^\rho = \kappa = [1-\mu_0\delta]/[(1-\mu_0)\delta]$, we have that

$$\begin{aligned} V^{1,2} &= \overline{V}(\hat{\beta}) = \mu_0 + \hat{a}\hat{\beta}^\rho \quad \text{and} \\ \overline{V}(\hat{\beta}/\lambda_1) &= \mu_0 + \hat{a}(\hat{\beta}/\lambda_1)^\rho = V^{1,2} - \left[V^{1,2} - \mu_0\right]\left[1 - \frac{1}{\lambda_1^\rho}\right] \\ &= V^{1,2} - \left[V^{1,2} - \mu_0\right]\frac{1-\delta}{1-\mu_0\delta}, \end{aligned}$$

which establishes the lower bound on $\overline{V}(\underline{\beta})$; the upper bound is proved similarly. $\square$

Let

$$V^{1,2}(\pi_B, 1) = \lim_{\delta\uparrow 1} V^{1,2}(\pi_B, \delta).$$

One can check that

$$V^{1,2}(\pi_B, 1) = V^M(\pi^*, \pi_B) = \mu_0 + (1-\mu_0)\frac{\log\left[\frac{1-\pi_B}{1-\pi^*}\right]}{\log\left[\frac{\pi^*(1-\pi_B)}{(1-\pi^*))\pi_B}\right]},$$

which we will return to interpret at the end of the Appendix.

Proof of Proposition 4:

Step 1: We first prove that for any $\epsilon \in (0, 1/2)$ there exists $\underline{\delta} \in (0,1)$ such that $\overline{V}(\beta, \delta) \leq V^{1,2}(\pi_B, 1) + \epsilon$ for all $\beta \in [\epsilon, 1-\epsilon]$ and $\delta \geq \underline{\delta}$. By Lemma 11, for any $\hat{\epsilon} \in (0, \epsilon/2]$ and any $\delta \geq \hat{\delta}(\hat{\epsilon})$, $\overline{V}(1-\hat{\epsilon}, \delta) - \overline{V}(\hat{\epsilon}, \delta) < \hat{\epsilon}$. Since $\bar{\pi}(1/2) < 2\pi^*$ and $\bar{\pi}$ is convex, $\bar{\pi}(\hat{\epsilon}) \leq (1-2\hat{\epsilon})\bar{\pi}(0) + 2\hat{\epsilon}\bar{\pi}(1/2) < \pi^* + 2\pi^*\hat{\epsilon}$. Therefore,

$$\pi^* \leq \pi(\beta) \leq \bar{\pi}(\beta) < \pi^* + 2\pi^*\hat{\epsilon} \quad \text{for all } \beta \in [0, \hat{\epsilon}].$$

By Lemma 10,

$$V^\pi(\beta, \delta) \leq V^M(\pi^* + 2\pi^*\hat{\epsilon}, \pi_B) + O([1-\delta]^2)$$

for all $\beta \in [0, \hat{\epsilon}]$. By Lemma 8, we get

$$\begin{aligned} \overline{V}(1-\hat{\epsilon}, \delta) - V^{1,2}(\pi_B, 1) &= [\overline{V}(1-\hat{\epsilon}, \delta) - \overline{V}(\hat{\epsilon}, \delta)] + [\overline{V}(\hat{\epsilon}, \delta) - V^{1,2}(\pi_B, 1)] \\ &< \hat{\epsilon} + V^M(\pi^* + 2\pi^*\hat{\epsilon}, \pi_B) - V^M(\pi^*, \pi_B) + O([1-\delta]^2). \end{aligned}$$

Since V^M is continuous, we can choose $\underline{\delta} \leq \hat{\delta}(\hat{\epsilon})$ sufficiently small so that the right-hand side is less than ϵ for all $\delta \geq \underline{\delta}$. This concludes the proof of Step 1.

Fix $0 < \epsilon < 1 - \beta_d$. By contradiction, assume that there exists a sequence $\delta_j \to 1$ such that for each δ_j there is a $\beta \in [0, 1-\epsilon]$ such that $\pi(\beta) > \pi^* + \epsilon$. Let $\beta_0 = 1-\epsilon$. Since π is monotone, this implies that $\pi(\beta_0) > \pi^* + \epsilon$ for all δ_j.

Step 2: Pick a target $\pi^{n\times 1} \in (\pi^*, 1)$ close to 1. How $\pi^{n\times 1}$ is selected is explained in Step 3. To simplify notation below, let $\Delta_j = (1-\delta_j)/\delta_j$. We now find $\bar{\delta} < 1$, $\eta > 0$ and $K \in \mathbb{N}$ that depend on $\pi^{n\times 1}$, and an increasing sequence $\{\beta_k\}$ with the property that for any $\delta_j \in [\bar{\delta}, 1)$ and $m = 1, 2, \ldots,$

$$\overline{V}(\beta_{mK}, \delta_j) \leq V^M(\pi^*) + \epsilon + mK\Delta_j \quad \text{and} \quad \pi(\beta_{mK}) \geq \pi(\beta_0) + m\eta.$$

The sequence stops at $m = M$ when $\pi(\beta_{MK-1}) > \pi^{n\times 1}$.

The function $V^M(\pi_0)$ is convex and strictly increasing. Therefore, for all $\pi_0 \geq \pi^*$,

$$V^M(\pi_0) \geq V^M(\pi^*) + \dot{V}^M(\pi^*)(\pi_0 - \pi^*) \quad \text{where} \quad \dot{V}^M(\pi^*) = \frac{dV^M}{d\pi_0}(\pi^*).$$

Let

$$\hat{\epsilon} = \min\left\{\epsilon, \frac{1}{3}\dot{V}^M(\pi^*)\epsilon\right\}.$$

By Step 1, there exists $\hat{\delta} < 1$ such that $\overline{V}(\beta, \delta) \leq V^{1,2}(\pi_B, 1) + \hat{\epsilon}$ for all $\beta \in [\hat{\epsilon}, 1-\hat{\epsilon}]$ and $\delta \in [\hat{\delta}, 1)$. In particular, $\overline{V}(\beta_0, \delta) \leq V^{1,2}(\pi_B, 1) + \hat{\epsilon}$ for all $\delta \in [\hat{\delta}, 1)$.

Fix $\delta_j \in [\hat{\delta}, 1)$. Starting at β_0, for $k \geq 0$, sequentially define

$$\beta_{k+1} = \beta_{L,\ell}(\beta_k, \pi(\beta_k)) \quad \text{and} \quad \beta_{k,H} = \beta_{H,\ell}(\beta_k, \pi(\beta_k)) = \frac{\beta_0}{\beta_0 + (1-\beta_0)L_k}$$
$$\text{where} \quad L_k = \left[\frac{1}{(1-\pi_B)^k}(1-\pi(\beta_0))(1-\pi(\beta_1))\cdots(1-\pi(\beta_{k-1}))\right]\frac{\pi(\beta_k)}{\pi_B}.$$

Also define $\pi_k = \pi(\beta_k)$. These concepts can be understood as follows. Suppose for $k+1$ times in a row the state is ℓ. Then, β_{k+1} is the reputation that would be obtained (from β_0) after the sender recommends L every time, and $\beta_{k,H}$ is the reputation that would be obtained after the sender recommends L for k times and H once. Clearly $\beta_k > \beta_{k-1}$ and $\beta_{k,H} > \beta_{k-1,H}$ for all k. Since $\pi_B < \pi^* < \pi_k \leq 1$ for all k,

$$L_k \leq \left[\frac{1-\pi^*}{1-\pi_B}\right]^k \frac{1}{\pi_B} \equiv \bar{L}_k.$$

Let K be the smallest integer such that $\bar{L}_{K-1} \leq 1$. Then $\beta_{K-1,H} \geq \beta_0$. Consider the initial posterior β_{K-1} and the associated interval $[\beta_H, \beta_L]$, where $\beta_H = \beta_{H,\ell}(\beta_{K-1}, \pi_{K-1}) = \beta_{K-1,H}$ and $\beta_L = \beta_{L,\ell}(\beta_{K-1}, \pi_{K-1}) = \beta_K$. By right promise keeping,

$$\overline{V}(\beta_{k+1}, \delta_j) - \overline{V}(\beta_k, \delta_j) < \overline{V}(\beta_{k+1}, \delta_j) - \overline{V}(\beta_{k,H}, \delta_j) = \frac{1-\delta_j}{\delta_j} = \Delta_j.$$

Hence

$$\overline{V}(\beta_k, \delta_j) \leq \overline{V}(\beta_0, \delta_j) + k\Delta_j \leq V^M(\pi^*) + \hat{\epsilon} + k\Delta_j \quad \text{for } k = K-1, K.$$

Since $\beta_{K-1} \geq \beta_0 \geq \beta_d$, $[\beta_{K-1}, \beta_L]$ does not intersect Region 1. Therefore,

$$V^M(\pi_{K-1}) \geq V^M(\pi_0) > V^M(\pi^*) + \dot{V}^M(\pi^*)\epsilon > V^M(\pi^*) + 3\hat{\epsilon}.$$

Let $\bar{\delta} = \max\{\hat{\delta}, K/(K+\hat{\epsilon})\}$, $E^{n\times 1} = \max\{E(\pi) \mid \pi \in [\pi^*, \pi^{n\times 1}]\}$, and $\eta = \hat{\epsilon}/E^{n\times 1}$. Then $K\Delta_j + E^{n\times 1}\eta \leq 2\hat{\epsilon}$ for all $\delta_j \in [\bar{\delta}, 1)$. Let $\delta_j \in [\bar{\delta}, 1)$. Then

$$\begin{aligned} \overline{V}(\beta_{K-1}, \delta) &\leq V^M(\pi_{K-1}) - 2\hat{\epsilon} + (K-1)\Delta_j \\ &\leq V^M(\pi_{K-1}) - (1-\delta) - E^{n\times 1}\eta. \end{aligned} \tag{$*$}$$

If $\pi_{K-1} > \pi^{n\times 1}$ stop and make $M = 1$. Otherwise, $E(\pi_{K-1}) \leq E^{n\times 1}$ and by Corollary 2 it must be that $\eta \leq \pi(\beta_L) - \pi(\beta_H) = \pi(\beta_K) - \pi(\beta_{K-1,H}) \leq \pi(\beta_K) - \pi(\beta_0)$, or $\pi_K \geq \pi_0 + \eta$.

We repeat this process again starting at β_K instead of β_0. The definition of K implies again that $\beta_{2K-1,H} \geq \beta_K$. By a similar argument as above, we have that

$$\begin{aligned} &\overline{V}(\beta_k, \delta_j) \leq \overline{V}(\beta_0, \delta_j) + k\Delta_j \leq V^M(\pi^*) + \hat{\epsilon} + k\Delta \quad \text{for } k = 2K-1, 2K \\ &\overline{V}(\beta_{2K-1}, \delta_j) \leq V^M(\pi_{2K-1}) - (1-\delta_j) - E^{n\times 1}\eta. \end{aligned}$$

If $\pi_{2K-1} > \pi^{n\times 1}$, stop and make $M = 2$. Otherwise, Corollary 2 again implies that $\pi_{2K} \geq \pi_K + \eta$. And so on. This concludes the proof of Step 2.

At the end of Step 2, for any $\delta_j \in [\bar{\delta}, 1)$ we stop at a posterior $\beta^{n\times 1} \equiv \beta_{MK}$ such that $\pi(\beta^{n\times 1}) \geq \pi(\beta_{MK-1}) > \pi^{n\times 1}$ Most importantly, though $\pi(\beta)$ changes with δ_j, Step 2 is guaranteed to stop in at most $\bar{M}$ rounds for any $\delta_j \in [\bar{\delta}, 1)$, where

$$\bar{M} = \frac{(\pi^{n\times 1} - \pi^*)E^{n\times 1}}{\hat{\epsilon}}.$$

Step 3: It is time to choose $\pi^{n\times 1}$. Let

$$V^{n\times 1} = \frac{2}{3} + \frac{1}{3}V^M(\pi^*).$$

Consider the $n \times 1$ strategy that always recommends H when $\theta = h$, and along the periods when $\theta = \ell$ it recommends the cycle $LHH\cdots HLHH\cdots HLHH\cdots HLHH\cdots$ of one L followed by n H's. We will show that when the sender follows this strategy, the receiver accepts all his recommendations and therefore the sender attains an expected discounted payoff arbitrarily close to

$$\mu_0 + (1-\mu_0)\frac{n}{n+1} = \frac{\mu_0 + n}{1+n}$$

as $\delta_j \to 1$. Let n be the smallest integer such that $[\mu_0 + n]/[1 + n] > V^{n\times 1}$. We want to choose $\pi^{n\times 1}$ close enough to 1 so that when the sender follows the $n \times 1$ strategy starting at $\beta^{n\times 1}$, the posterior always remains above $\beta^{n\times 1}$ along the stochastic path. Starting at $\tilde{\beta}_0 \geq \beta^{n\times 1}$ let us follow the posterior during one cycle of the $n\times 1$ strategy. The posterior does not change in periods where $\theta = h$. Let $\tilde{\beta}_1$ be the posterior after the first period where $\theta = \ell$ and the sender recommends H. Let $\tilde{\beta}_k$ be the posterior after k periods with $\theta = \ell$ where the sender has recommended L once and then H for $k-1$ periods, $k = 2, \dots, n+1$. Obviously $\tilde{\beta}_1 > \tilde{\beta}_0$ and $\tilde{\beta}_{n+1} < \tilde{\beta}_n < \cdots < \tilde{\beta}_2 < \tilde{\beta}_1$. To ensure that the posterior remains above $\tilde{\beta}_0$ it is enough to verify that $\tilde{\beta}_{n+1} \geq \tilde{\beta}_0$. We have that

$$\tilde{\beta}_{n+1} = \frac{\tilde{\beta}_0}{\tilde{\beta}_0 + (1-\tilde{\beta}_0)L} \quad \text{where} \quad L = \frac{(1-\tilde{\pi}_0)\tilde{\pi}_1 \cdots \tilde{\pi}_n}{(1-\pi_B)\pi_B^n}$$

and $\tilde{\pi}_k = \pi(\tilde{\beta}_k)$, $k = 0, \dots, n+1$. Let

$$\pi^{n\times 1} = 1 - (1-\pi_B)\pi_B^n.$$

Then

$$L < \frac{1-\tilde{\pi}_0}{(1-\pi_B)\pi_B^n} \leq \frac{1-\pi^{n\times 1}}{(1-\pi_B)\pi_B^n} = 1,$$

and $\tilde{\beta}_{n+1} \geq \tilde{\beta}_0$ as desired. Since $\tilde{\beta}_0 \geq \beta^{n\times 1} \equiv \beta_{MK} > \beta_0 = 1-\epsilon > \beta_d$, when the sender follows the $n \times 1$ strategy, the posterior remains above β_d in every period, and by the Remark following Lemma 8, the receiver accepts all the sender's recommendations, as we claimed above.

Finally, we show that this leads to a contradiction. Let $\delta^{n\times 1} \in [\bar{\delta}, 1) \cap \{\delta_j\}$ be such that

$$\overline{V}(\beta^{n\times 1}, \delta^{n\times 1}) \leq V^M(\pi^*) + \epsilon + (\bar{M}K)\frac{1-\delta^{n\times 1}}{\delta^{n\times 1}} \leq \frac{1}{3} + \frac{2}{3}V^M(\pi^*). \tag{$*$}$$

Note that $V^{n\times 1} - \overline{V}(\beta^{n\times 1}, \delta^{n\times 1}) \geq [1 - V^M(\pi^*)]/3 > 0$. After arriving at $\beta^{n\times 1}$ in Step 2, the sender can follow the $n \times 1$ strategy forever because the posterior never drops below $\pi^{n\times 1}$, and hence can attain a continuation value larger than $V^{n\times 1}$. So $\overline{V}(\beta^{n\times 1}, \delta^{n\times 1}) \geq V^{n\times 1}$, which contadicts $(*)$ above. Therefore, there exists $\underline{\delta} \in (0,1)$ such that for all $\delta \in [\underline{\delta}, 1)$, $\pi(\beta) \leq \pi^* + \epsilon$ for all $\beta \in [0, 1-\epsilon]$.

Step 4: We finally prove that for any $\epsilon > 0$ there exists $\underline{\delta} \in (0,1)$ such that $\overline{V}(\beta, \delta) \geq V^{**}(\pi_B) - \epsilon$ for all $\beta \in [\epsilon, 1-\epsilon]$ and $\delta \geq \underline{\delta}$. By previous argument and Lemma 11, for any $\hat{\epsilon} \in (0, \epsilon/2]$ there exists $\hat{\delta} \in (0,1)$ such that $\overline{V}(1-\hat{\epsilon}, \delta) - \overline{V}(\hat{\epsilon}, \delta) \leq \hat{\epsilon}$ and $\pi^* \leq \pi(\beta) \leq \pi^* + \hat{\epsilon}$ for all $\beta \in [\hat{\epsilon}, 1-\hat{\epsilon}]$ and $\delta \in [\hat{\delta}, 1)$. Choose $\hat{\epsilon} \leq [\pi^* - \pi_B]/2$. Since $\bar{\pi}$ is convex, $\bar{\pi}(\beta) \geq \bar{\pi}(0) + \bar{\pi}'(0)\beta = \pi^* + (\pi^* - \pi_B)\beta$. Therefore $\bar{\pi}(\beta) > \pi^* + \hat{\epsilon}$ for all $\beta \in [1/2, 1]$. This implies $\pi(\beta) < \bar{\pi}(\beta)$ for all $\beta \in [1/2, 1-\hat{\epsilon}]$ and $[1/2, 1-\hat{\epsilon}]$

does not intersect Region 1. Pick any $\beta_0 > 1/2$ such that $[\beta_H, \beta_L] \subseteq [1/2, 1-\hat{\epsilon}]$, where $\pi_0 = \pi(\beta_0)$, $\beta_H = \beta_{H,\ell}(\beta_0, \pi_0)$ and $\beta_L = \beta_{L,\ell}(\beta_0, \pi_0)$. By Lemmas 8 and 10,

$$\begin{aligned}\overline{V}(\beta_0, \delta) = V^\pi(\beta_0, \delta) &\geq V^M(\pi_0) + O([1-\delta]^2) - E(\pi^*)\hat{\epsilon} \\ &\geq V^M(\pi^*) + O([1-\delta]^2) - E(\pi^*)\hat{\epsilon}.\end{aligned}$$

We can choose $\hat{\epsilon}$ and $\underline{\delta} \leq \hat{\delta}(\hat{\epsilon})$ such that the right hand side is greater or equal to $V^M(\pi^*) - \epsilon/2 = V^{**}(\pi_B) - \epsilon/2$ for all $\delta \in [\underline{\delta}, 1)$. Therefore, for any $\beta \in [\hat{\epsilon}, 1-\hat{\epsilon}] \supset [\epsilon, 1-\epsilon]$ and any $\delta \in [\underline{\delta}, 1)$,

$$\overline{V}(\beta, \delta) \geq \overline{V}(\hat{\epsilon}, \delta) \geq \overline{V}(\beta_0, \delta) - \hat{\epsilon} \geq V^{**}(\pi_B) - \epsilon. \qquad \square$$

Proof of Lemma 14: Let $\bar{\gamma} = \epsilon^4$. We first show that for all $\pi_B \in [\pi^* - \bar{\gamma}, \pi^*)$ and for all $\beta \in [0, 1-\epsilon^2]$, $\pi^* \leq \pi(\beta) \leq \pi^* + \epsilon^2$. Let $\gamma = \pi^* - \pi_B$ and $\beta \in [0, 1-\epsilon^2]$. Then

$$\bar{\pi}(\beta) = \frac{\pi^* - \beta\pi_B}{1-\beta} = \pi^* + \frac{\beta}{1-\beta}\gamma \leq \pi^* + \frac{\beta}{1-\beta}\epsilon^4 \leq \pi^* + \epsilon^2.$$

Since $\pi^* \leq \pi(\beta) \leq \bar{\pi}(\beta)$, $\pi^* \leq \pi(\beta) \leq \pi^* + \epsilon^2$, as claimed. Moreover, since $\pi^* + \epsilon^2 < 1$, this also implies that $R_3 \subset (1-\epsilon^2, 1]$ and $R_1 \cup R_2 \supset [0, 1-\epsilon^2]$, where R_i denotes Region i, $i = 1, 2, 3$.

Recall that

$$\lambda_1 = \frac{(1-\pi_B)(1-\mu_0)}{1-2\mu_0} = \frac{1-\pi_B}{1-\pi^*} = 1 + \frac{\gamma}{1-\pi^*} < 1 + 2\epsilon^4.$$

Let $\beta^0 \in [0, \lambda_1(1-\epsilon)]$, and recursively define $\beta^{k+1} = \beta^k_{L,\ell}$. Since $\pi(\beta^k) \leq \bar{\pi}(\beta^k)$, we have that $\beta^{k+1} \leq \lambda_1\beta^k$. Note that $\log((1+\epsilon)/\lambda_1) > \epsilon/2$ for $\epsilon > 0$ small, and $\log(1+2\epsilon^4) < 2\epsilon^4$. Therefore,

$$\beta^k \leq \lambda_1^k\beta^0 < (1+2\epsilon^4)\lambda_1(1-\epsilon) < 1-\epsilon^2$$

for all $k \leq \bar{k} \equiv 1/[4\epsilon^3]$. Since $\overline{V}(\beta^0) \geq 0$ and $\overline{V}(\beta^{\bar{k}}) \leq 1$, this implies that on average

$$\Delta^k = \overline{V}(\beta^{k+1}) - \overline{V}(\beta^k) \leq 1/\bar{k} = 4\epsilon^3 < \epsilon^2.$$

Therefore, there exists $k \leq \bar{k}$ such that $\Delta^k < \epsilon^2$. For that k, consider the right promise keeping condition for β^k:

$$\overline{V}(\beta^k) = \mu_0(1-\delta) + \frac{\delta}{4}[\mu_0\overline{V}(\beta^k) + (1-\mu_0)\overline{V}(\beta^{k+1}) < \mu_0(1-\delta) + \delta\overline{V}(\beta^k) + (1-\mu_0)\delta\epsilon^2.$$

That is,

$$\overline{V}(\beta^k) < \mu_0 + \frac{\delta}{1-\delta}(1-\mu_0)\epsilon^2 < \mu_0 + \epsilon.$$

This implies that $\overline{V}(\beta^0) < \mu_0 + \epsilon$ for all $\beta^0 \in [0, \lambda_1(1-\epsilon)]$. We prove below that $V^{1,2}(\pi_B, \delta)$ is increasing in π_B and $\mu_0 + \epsilon < V^{1,2}(\pi^* - \bar{\gamma}, \delta)$ (for $\epsilon > 0$ sufficiently small). Therefore, if Region 1b is empty, Lemma 13 implies that $R_1 \supset [0, 1-\epsilon]$. Recall that

$$V^{1,2}(\pi_B, \delta) = \mu_0 + \frac{1-\delta}{\delta[\kappa - \kappa^{\ell_0/\ell_1}]},$$

where $\kappa > 1$ is a constant (it does not depend on π_B), and

$$\ell_0 = \log(\lambda_0) = \log\left(1 - \frac{\gamma}{\pi^*}\right) \quad \text{and} \quad \ell_1 = \log(\lambda_1) = \log\left(1 + \frac{\gamma}{1-\pi^*}\right).$$

Since $\ell_0/\ell_1 < 0$ and

$$\frac{d(\ell_0/\ell_1)}{d\gamma} < 0,$$

$V^{1,2}(\pi_B, \delta)$ is increasing in π_B and $\mu_0 + \epsilon < V^{1,2}(\pi_B, \delta)$ for all $\pi_B \in [\pi^* - \bar{\gamma}, \pi^*)$, provided that $0 < \epsilon < V^{1,2}(\pi^* - \bar{\gamma}, \delta)$. □

Remark: It is now easier to see why at very high values of δ, the value of $\overline{V}(\underline{\beta})$ is well approximated by $V^M(\pi^*, \pi_B)$, as asserted at the end of the proof of Lemma 13. Doing "reputation maintenance" at any β in Region 2 yields a stream of actions H and L by the receiver. If the sender were ideally patient, he would care only about the proportion of times he induced H and L, respectively. Pretending that π is perfectly flat in the reputation-maintenance region of β gives us a simple expression for $\overline{V}(\beta)$ in terms of the ratio π/π_B (see V^M defined in Section 6.2).

We now know that as δ approaches 1, Region 1 vanishes asymptotically. Thus, doing reputation maintenance at $\underline{\beta}$, where by definition $\pi(\underline{\beta}) = \bar{\pi}(\underline{\beta})$, we see that as δ approaches 1, $\pi(\underline{\beta})$ tends to π^* (the KG commitment ideal, also the vertical intercept of the $\bar{\pi}$ function). Now for δ close to 1, the sender is close to ideally patient, and π is virtually constant in the (vanishingly short) interval of reputation maintenance around $\underline{\beta}$. So as Region 2 begins, $\overline{V}$ asymptotically takes the value $V^M(\pi^*, \pi_B)$. Indeed, this could be called the value of the game, as $\overline{V}$ is virtually flat except at values of β extremely close to 0 or 1.

References

Abreu, Dilip and Faruk Gul (2000), "Bargaining and Reputation," *Econometrica* **68**(1), 85–117.

Abreu, Dilip, Paul Milgrom and David Pearce (1991), "Information and Timing in Repeated Partnerships," *Econometrica* **59**(6), 1713–1733.

Abreu, Dilip, David Pearce and Ennio Stacchetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* **58**(5), 1041–1064.

Abreu, Dilip and David Pearce (2007), "Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts," *Econometrica* **75**(3), 653–710.

Aumann, Robert and Sergiu Hart (2003), "Long cheap talk," *Econometrica* **71**(6), 1619–1660.

Aumann, Robert and Michael Maschler (1995), "Repeated Games with Incomplete Information," MIT Press, (Chapter 1).

Backus, David and Edward Driffill (1985), "Inflation and Reputation," American Economic Review **75**, 530–538.

Bénabou, Roland, and Guy Laroque (1992), "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics* **107**, 921–958.

Benoît, Jean-Pierre and Juan Dubra (2011), "Apparent Overconfidence," *Econometrica* **79**(5), 1591–1625.

Bergemann, Dirk and Stephen Morris (2016), "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics* **11**, 487–522.

Bergemann, Dirk and Stephen Morris (2019), "Information Design: A Unified Perspective," *Journal of Economic Literature* **57**(1), 44–95.

Best, James, and Daniel Quigley (2017), "Persuasion for the long run." Economic Papers 2016-w12, Economics Group, Nuffield College, University of Oxford.

Borovicka, Jaroslav and John Stachurski (2020), "Necessary and Sufficient Conditions for Existence and Uniqueness of Recursive Utilities", *Journal of Finance* **75**(3), 1457–1493.

Crawford, Vincent and Joel Sobel (1982), "Strategic Information Transmission," *Econometrica* **50**(6), 1431–1451.

Cripps, Martin W., George J. Mailath and Larry Samuelson (2004), "Imperfect Monitoring and Impermanent Reputations", *Econometrica*, **72**(2), 407–432.

Ely, Jeffrey (2017), “Beeps,” *American Economic Review* **107**(1), 31–53.

Ely, Jeffrey and Juuso Valimäki (2003), “Bad Reputation,” *Quarterly Journal of Economics* **118**(3), 785–814.

Fanning, Jack (2016), “Reputational Bargaining and Deadlines,” *Econometrica* **84**(3), 1131–1179.

Fernandes, Ana and Christopher Phelan (2000), “A Recursive Formulation for Repeated Agency with History Dependence,” *Journal of Economic Theory* **91**, 223–247.

Fudenberg, Drew and David Levine (1989), “Reputation and Equilibrium Selection in Games with a Patient Player,” *Econometrica* **57**(4), 759–778.

Fudenberg, Drew and David Levine (1992), “Maintaining a Reputation when Strategies are Imperfectly Observed,” *Review of Economic Studies* **59**, 561–579.

Fudenberg, Drew and David Levine (1994), “Efficiency and Observability with Long-run and Short-run Players,” *Journal of Economic Theory* **62**(1), 103–135.

Fudenberg, Drew, David Kreps and Eric Maskin (1990), “Repeated Games with Long-run and Short-run Players,” *Review of Economic Studies* **57**, 555–573.

Fudenberg, Drew, David Levine and Eric Maskin (1994), “The Folk Theorem with Imperfect Public Information,” *Econometrica* **62**(5), 997–1039.

Golosov, Mikhail, Vasiliki Skreta, Aleh Tsyvinski and Andrea Wilson (2014), “Dynamic strategic information transmission,” *Journal of Economic Theory*, **151**, 304–341.

Golosov, Mikhail, Aleh Tsyvinski and Ivan Werning (2007), “New Dynamic Public Finance: A User’s Guide,” NBER Chapters, in: NBER Macroeconomics Annual 2006, **21**, 317–388.

Goltsman, Maria, Johannes Hörner, Gregory Pavlov and Francesco Squintani (2009), “Mediation, Arbitration and Negotiation,” *Journal of Economic Theory*, **144**(4), 1397-1420.

Gorno, Leandro and Felipe S. Iachan (2020), “Competitive Real Options Under Private Information,” *Journal of Economic Theory*, **185**, 104945.

Hörner, Johannes, Satoru Takahashi and Nicolas Vieille (2015), “Truthful Equilibria in Dynamic Bayesian Games,” *Econometrica* **83**(5), 1795–1848.

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

Kamenica, Emir (2019), “Bayesian Persuasion and Information Design,” *Annual Review of Economics* **11**, 249–272.

Kandori, Michihiro and Ichiro Obara (2006), "Efficiency in Repeated Games Revisited: The Role of Private Strategies," *Econometrica* **74**(2), 499–519.

Koessler, Frederic and Françoise Forges (2008a): "Multistage Communication with and without Verifiable Types," *International Game Theory Review* **10**, 145–164.

Koessler, Frederic and Françoise Forges (2008b): "Transmission strategique de l'infor- mation et certification," *Annales d'Economie et Statistiques* **89**, 3–61.

Kreps, David and Bob Wilson (1982), "Reputation and Imperfect Information," *Journal of Economic Theory* **27**(2), 253–279.

Kreps, David, Paul Milgrom, John Roberts and Robert Wilson (1982), "Rational cooperation in the finitely repeated prisoners' dilemma," *Journal of Economic Theory* **27**(2), 245–252.

Krishna, Vijay and John Morgan (2004), "The art of conversation: eliciting information from experts through multi-stage communication," *Journal Economic Theory* **117**(2), 147–179.

Mailath, Georges and Larry Samuelson (2001), "Who Wants a Good Reputation," *Review of Economic Studies* **68**, 415–441.

Mailath, Georges and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press.

Margaria, Chiara and Alexander Smolin (2015), "Dynamic Communication with Biased Senders," *Games and Economic Behavior* **110**, 330–339.

Mathevet, Laurent, Jacopo Perego and Ina Taneva (2019), "On Information Design in Games", *Journal of Political Economy* **128**(4), 1370–1404.

Milgrom, Paul and John Roberts (1982), "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory* **27**(2), 280–312.

Morris, Stephen (2001), "Political Correctness," *Journal of Political Economy* **109** (2), 231–265.

Ottaviani, Marco and Peter Sorensen (2006a), "Professional advice," *Journal of Economic Theory* **126**(1), 120–142.

Ottaviani, Marco and Peter Sorensen (2006b), "Reputational cheap talk," *RAND Journal of Economics* **3**(1), 155–175.

Pei, Harry (2021), "Repeated Communication with Private Lying Costs," Working Paper.

Phelan, Christopher (2006), "Public trust and government betrayal," *Journal of Economic Theory* **130**, 27–43.

Phelan, Christopher and Ennio Stacchetti (2001), "Sequential Equilibria in a Ramsey Tax Model," *Econometrica* **69**(6), 1491–1518.

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

Rayo, Luis and Ilya Segal (2010), "Optimal Information Disclosure," *Journal of Political Economy* **118**(5), 949–987.

Renner, Philipp and Simon Scheidegger (2020), "Machine Learning for Dynamic Incentive Problems," *Working Paper*.

Shamsie, Jamal (2003), "The Context of Dominance: An Industry-driven Framework for Exploiting Reputation", *Strategic Management*, **24**(3), 199–215.

Sobel Joel (1985),"A theory of credibility," *Review of Economic Studies* **52**(4), 557–573.

Taneva, Ina (2019), "Information Design", *American Economic Journal: Microeconomics* **11**(4), 151–185.

Wolitzky, Alexander (2012), "Reputational Bargaining With Minimal Knowledge of Rationality", *Econometrica* **80**(5), 2047–2087.

A. The Basic Model: Repeated Cheap Talk

This Section considers an infinitely repeated Cheap Talk game between a long-lived sender and a sequence of short-lived receivers, one in each period. We consider two alternative informational conditions concerning what is observable about histories of play, differing by whether or not the sender's mixed strategies are observable. We derive the sender's optimal equilibrium payoff under both conditions.

A.1 The Stage Game

There is a finite set of states Θ and a finite set of actions A. The receiver's payoff $u(\theta, a)$ and the sender's payoff $v(\theta, a)$ depend on the receiver's action $a \in A$ and state $\theta \in \Theta$. These payoff functions are extended to mixed actions of the receiver, $\alpha \in \Delta(A)$, by taking expectations and are denoted $u(\theta, \alpha)$ and $v(\theta, \alpha)$. The state is drawn randomly according to $\mu_0 \in \Delta(\Theta)$, where $\Delta(\Theta)$ denotes the set of probability distributions over Θ. After privately observing the state, the sender sends a message from finite set M to the receiver. A strategy for the sender is a mapping $\pi : \Theta \to \Delta(M)$. For each $\theta \in \Theta$, let $\pi(m|\theta)$ be the probability that the sender sends message m when he observes state θ.

Given sender's strategy π and received message m, the receiver formulates posterior belief

$$\mu_m^\pi(\theta) = \frac{\pi(m|\theta)\mu_0(\theta)}{\pi(m)} \quad \text{where} \quad \pi(m) = \sum_{\theta' \in \Theta} \pi(m|\theta')\mu_0(\theta').$$

Then, she chooses an optimal action $a_m \in A^*(\mu_m^\pi)$ where

$$A^*(\mu) = \operatorname*{argmax}_{a \in A} \sum_{\theta \in \Theta} \mu(\theta) u(\theta, a)$$

is the set of receiver's optimal actions at posterior belief $\mu \in \Delta(\Theta)$.

A.2 The Repeated Game

States $\{\theta_t\}$ are drawn i.i.d. with distribution μ_0 in each period. At the end of period $t-1$, the state and the sender's message in that period, (θ_{t-1}, m_{t-1}), become commonly known to all (future) players. Under Ideal Observational Conditions (IOC), the strategy chosen by the sender, π_{t-1}, is *also* publicly observed at the end of period $t-1$. Therefore, the period t receiver is assumed to know the mixed strategies used by the sender, the messages he sent, and the true states of nature, in every one of the first $t-1$ periods. Given such a history and the period t message, the receiver takes an action a_t. The receiver and the sender then get their payoffs $(u(\theta_t, a_t), v(\theta_t, a_t))$.

Just as information design is vulnerable to the observation that it is hard, as a practical matter, for a sender to commit legally and costlessly to a particular information protocol, IOC is open to attack. How often is it really possible to observe not just what someone did, but what probabilities he used to determine his choice? Accordingly, after studying IOC, we weaken the information at the disposal of a receiver. Under Normal Observational Conditions (NOC), the period t receiver is assumed instead to know only the messages sent by the sender and the true state, in every one of the first $t-1$ periods.

Let $\delta \in (0,1)$ be the sender's discount factor and let $V(\delta)$ denote the set of PBE values for the sender. A public randomization device ensures that $V(\delta)$ is a convex set: at the begining of every period, before the state is realized, the players publicly observe the outcome of a draw from a uniform distribution in $[0,1]$. Using the self-generation result of APS (1990), one can also show that $V(\delta)$ is a compact set. Let

$$\overline{v}(\delta) = \max \; V(\delta).$$

Our results will compare two bounds. The first one is the standard information-design value that a sender with commitment can attain:

$$\begin{aligned} v^* = \max_{\pi,(a_m)} \quad & \sum_{\theta\in\Theta}\sum_{m\in M} \mu_0(\theta)\pi(m|\theta)v(\theta,a_m) \\ \text{s.t.} \quad & \pi(\cdot|\theta) \in \Delta(M) \text{ for all } \theta\in\Theta, \\ & \text{and } a_m \in A^*(\mu_m^\pi) \text{ for all } m \in M. \end{aligned}$$

Let $(\pi^*, \{a_m^*\})$ denote an optimal solution to this problem.[21]

The second bound is given by another static optimization problem,

$$\begin{aligned} v^N = \max_{\pi,(\alpha_m)} \quad & \sum_{\theta\in\Theta} \mu_0(\theta) \min_{m\in M}\{v(\theta,\alpha_m) \mid \pi(m|\theta)>0\} \\ \text{s.t.} \quad & \alpha_m \in \Delta(A) \text{ for all } m\in M, \quad \pi(\cdot|\theta)\in\Delta(M) \text{ for all } \theta\in\Theta \\ & \text{and } \alpha_m(a_m) = 0 \text{ for all } a_m \notin A^*(\mu_m^\pi), \end{aligned}$$

where the sender's payoff in every state is given by the worst action induced with positive probability in that state.[22] Thus, $v^N \le v^*$. Denote by $(\pi^N, \{\alpha_m^N\}_{m\in M})$ an optimal solution.

[21] Most of the information design literature is concerned with the joint optimal choice of a message space M and signal π. We will assume instead that M is exogenously given. Our results hold for an arbitrary message space M (whether it is optimal or not).

[22] Mixed actions do not change the value of v^N but are used in the proof of Proposition A.2.

Proposition A.1. *Under IOC, there exists* $\underline{\delta} < 1$ *such that* $\overline{v}(\delta) = v^*$ *for all* $\delta \in [\underline{\delta}, 1)$.

This proposition displays an PBE that lets the sender do exactly as well as he would if he could bind himself legally to use any information protocol. Here, intertemporal incentives take the place of legal enforcement. As a result, all of the "concave envelope" technology of KG and the linear programming of Bergemann and Morris (2016) are applicable. To find the sender's best PBE in the infinite horizon policy game, just solve the static problem.

The proof is straightforward and hence omitted. It is well known, for example from Crawford and Sobel (1982), that "babbling" is always an equilibrium of the stage game. Therefore, playing the babbling equilibrium in every period is a PBE for all δ. In the optimal PBE, the sender uses the information design optimal signal in every period on path, thereby achieving v^*, while "reversion to babbling forever" is used as punishment if the sender deviates in any period from his equilibrium strategy.

Proposition A.2. *Under NOC,* $\overline{v}(\delta) \leq v^N$ *for all* $\delta \in (0, 1)$ *and there is* $\underline{\delta} < 1$ *such that* $\overline{v}(\delta) = v^N$ *for all* $\delta \in [\underline{\delta}, 1)$.

Under NOC, randomization by the sender is expensive in a way it is not under IOC. This is an example of a phenomenon long understood by game theorists (see Fudenberg, Kreps and Maskin (1990, p. 562)). Consider an equilibrium in which the sender is supposed to randomize in the first period over two messages m_1 and m_2 after observing a particular state θ. If his myopic payoff $v(\theta, a_1)$ is higher than $v(\theta, a_2)$, where a_i is the receiver's best reply to the posterior induced by m_i, the only way to make the sender indifferent between the two, and hence willing to randomize, is to make his period-2 continuation value sufficiently lower when he says m_1 than when he says m_2. The need to use inefficient continuation values with positive probability bounds the sender's payoff strictly below the information-design value. The proof formalizes this intuition via a self-generation argument.

The inefficiency arises only if attaining the information-design value requires randomization over messages at a particular state. If, instead, it is deterministic (which means it is a partition of the state space, but not necessarily fully informative), then $v^N = v^*$ and there is no loss relative to information design in the best PBE. While randomizations can still be optimal in the repeated game, as a sacrifice in one state to induce a new action in another, they are not in two-state two-action games:

Corollary. *If* $\Theta = A = \{0, 1\}$, *then there is* $\underline{\delta} < 1$ *such that, for all* $\delta \in [\underline{\delta}, 1)$, *the sender optimal PBE is either the babbling or the fully informative PBE.*

We end this Section with two remarks:

Best and Quigley (2017) establish similar results in their Proposition 1 and Theorems 1 and 2. Proposition A.2 above is Proposition 1 in Best and Quigley (2017) (which they write in the space of random posteriors). Instead of assuming observable mixed strategies, their 'coin and cup' mechanism (in Theorem 2) supposes that there is a payoff-irrelevant random variable, whose realization is observable to the sender before he makes his period t choice, but observable to the receiver only after she has taken her period t action. That random variable reproduces the optimal information-design randomization. This sounds as though the sender's mixed strategy is observable, but there is a difference: In our case, the sender must be willing to follow his equilibrium randomization after observing the state, while, in theirs, she must be willing to play the corresponding message after observing the realization of the random variable.

Furthermore, there will typically be many PBEs in the repeated game and only the best one for the sender has any chance of matching the legal commitment value v^*.

A.3 Proof of Proposition A.2

Note that the stage game has a babbling equilibrium where the sender sends messages $m \in M$ with probability $\pi_B(m|\theta) = 1/|M|$ and the receiver chooses an action $a_B \in A^*(\mu_0)$. Let

$$v_B = \min_{a_B \in A^*(\mu_0)} \sum_{\theta \in \Theta} \mu_0(\theta) v(\theta, a_B)$$

be the worst babbling equilibrium value for the sender.

Proof: For any $\delta \in (0,1)$ consider an equilibrium that delivers the value $\overline{v}(\delta)$ to the sender. Let π be the sender's strategy in the first period, and α_m be the corresponding strategy for the receiver when she receives the message m. Denote by $w_{\theta,m}$ the sender's continuation value from period 2 onward when in the first period the state is θ and he sends the message m.

If $\pi(m_1|\theta) > 0$ and $\pi(m_2|\theta) > 0$ for two messages $m_1 \neq m_2$, it must be the case that the sender is indifferent between m_1 and m_2 when the state is θ. That is

$$(1-\delta)v(\theta, \alpha_{m_1}) + \delta w_{\theta,m_1} = (1-\delta)v(\theta, \alpha_{m_2}) + \delta w_{\theta,m_2}.$$

Let

$$\underline{v}(\theta) = \min\{v(\theta, \alpha_m) \mid \pi(m|\theta) > 0\}.$$

Since the equilibrium delivers the highest value to the sender, it must be that

$$w_{\theta,m} = \overline{v}(\delta) - \frac{1-\delta}{\delta}[v(\theta, \alpha_m) - \underline{v}(\theta)] \quad \text{for all} \quad m \in M \text{ with } \pi(m|\theta) > 0,$$

and the continuation value for the sender after observing the state θ (and before sending any messages) is $(1-\delta)\underline{v}(\theta) + \delta\overline{v}(\delta)$. Therefore

$$\overline{v}(\delta) = \sum_{\theta \in \Theta} \mu_0(\theta)[(1-\delta)\underline{v}(\theta) + \delta\overline{v}(\delta)],$$

so

$$\overline{v}(\delta) = \sum_{\theta\in\Theta} \mu_0(\theta)\underline{v}(\theta) \leq v^N.$$

To establish that the upper bound is attained for discount rates sufficiently high, we will show that the set $W = [v_B, v^N]$ is self-generating. If $v^N = v_B$ the PBE where the babbling equilibrium is played in every period has value v_B, and we are done. Therefore we now assume that $v^N > v_B$. For each $m \in M$, let

$$\tau^N(m) = \sum_{\theta\in\Theta} \mu_0(\theta)\pi^N(m|\theta).$$

Note that in the optimal solution $(\pi^N, \{\alpha^N_m\}_{m\in M})$ defined earlier, strategies α^N_m for signals m such that $\tau^N(m) = 0$ (i.e., that are never used) are arbitrarily specified. If $\tau^N(m) = 0$, we now define α^N_m by $\alpha^N_m(a_B) = 1$ and $\alpha^N_m(a) = 0$ for all $a \neq a_B$. For each $\theta \in \Theta$, let

$$\underline{v}^N(\theta) = \min\{v(\theta, \alpha^N_m) \mid \pi^N(m|\theta) > 0\}.$$

For each (θ, m) define the continuation values

$$w^N_{\theta,m} = \begin{cases} v_B & \text{if} \quad \pi^N(m|\theta) = 0 \\ v^N - \frac{1-\delta}{\delta}[v(\theta, \alpha^N_m) - \underline{v}^N(\theta)] & \text{if} \quad \pi^N(m|\theta) > 0. \end{cases}$$

Suppose the sender uses the signal π^N in the first period expecting the continuation values $\{w^N_{\theta,m}\}$, while the receiver follows the mixed strategy $\alpha^N = \{\alpha^N_m\}$ to (randomly) choose an action in the first period.

By definition, $w^N_{\theta,m} \leq v^N$ for all (θ, m). We now show that there exists $\underline{\delta} \in (0, 1)$ such that for all $\delta \in [\underline{\delta}, 1)$, $w^N_{\theta,m} \geq v_B$, and hence $w^N_{\theta,m} \in W$, for all (θ, m). Let

$$\Delta = \max\{v(\theta, a_1) - v(\theta, a_2) \mid a_1, a_2 \in A,\ \theta \in \Theta\}$$

and $\underline{\delta}$ be such that $[(1 - \underline{\delta})/\underline{\delta}]\Delta = v^N - v_B$ (if $\Delta = 0$, make $\underline{\delta} = 0$). Clearly, for any $\delta \in [\underline{\delta}, 1)$, $w^N_{\theta,m} \geq v_B$. Finally, note that

$$\sum_{\theta\in\Theta} \mu_0(\theta)\pi^N(m|\theta)[(1 - \delta)v(\theta, \alpha^N_m) + \delta w^N_{\theta,m}] = v^N.$$

This establishes that v^N is "generated" when the continuation values W are available (and $\delta \in [\underline{\delta}, 1)$). Using the babbling equilibrium, one can easily see that v_B is also generated with the continuation values in W (for any $\delta \in (0, 1)$). Since we assume the existence of a public randomization device, any value in $\text{co}\{v_B, v^N\} = W$ is also generated by continuation values W. Hence, W is self-generating. By APS, $W \subseteq V(\delta)$ for $\delta \in [\underline{\delta}, 1)$, which implies that $\overline{v}(\delta) = v^N$ for $\delta \in [\underline{\delta}, 1)$. □

B. The General Reputational Model

In the models of the preceding Section, the connection between the sender's behavior today and the receivers' expectations about his future behavior, can be quite arbitrary and unrealistic. For example, there are equilibria in which the sender "babbles" (sends uninformative random messages) for the first seventeen periods, but is informative and instantly trusted from period eighteen onward. In reality, agents' expectations are affected more systematically by the behavior of others. Introducing reputational types into the NOC model lets the sender invest or disinvest in his reputation for reliable communication. Understanding how this disciplines his behavior, in contrast to the usual appeal to legal enforcement, is, of course, the main subject of this paper.

The reputational model is an incomplete information version of the model from Section A, keeping the NOC but adding the possibility that the sender may be "behavioral." Suppose the sender can be either rational (type s_R), with preferences as in Section A.2, or a behavioral type that follows a fixed strategy in every period regardless of history. Let $\Pi_B \subseteq \{\hat{\pi} : \Theta \to \Delta(M)\}$ be the set of behavioral types, assumed to be finite, and let $\hat{\Pi} = \Pi_B \cup \{s_R\}$ be the set of all types.

Suppose Nature draws a prior β_0 uniformly from $\Delta(\hat{\Pi})$,[23] and then draws a sender's type, once and for all, according to β_0.

In each period, the sender communicates with the receiver about the state, causing her to update her belief and act. Since the sender's identity matters to interpret his message, the current receiver's belief about the sender's type is updated at the beginning of every period. This is done by updating the previous receiver's initial belief based on the message sent by the sender and the true state in that period. This material is formalized in Section 3.2.1.

Let $V(\beta; \delta)$ be the PBE value correspondence for the rational sender, given discount rate δ and the receiver's belief β. When δ is fixed, we will simply write $V(\beta)$.

The introduction of behavioral types with low probabilities has dramatic welfare consequences. If a patient sender has a rich set of behavioral types to imitate, he can achieve virtually the same payoff as his information-design value. This asymptotic efficiency applies to *all* PBE of the model (whereas *no* equilibrium of the NOC model of Section 2 comes close to that benchmark when optimal disclosure requires randomization). Proposition 3 states this welfare result precisely, as an application of Fudenberg and Levine (1992).

Concerns that asymptotic efficiency might be vacuous, because of nonexistence of equilibrium, led us to provide a series of results that prove existence of PBE for the entire class of games. The self-generation and algorithmic results of APS (1990) are extended to the reputational setting, and exploited to prove equilibrium existence. These

[23] Having Nature draw the prior at the beginning of the game is just a convenient device to study all the games beginning at different priors β_0 at once. The uniform distribution plays no role at all in the analysis.

techniques are used in Section 4 to explore qualitative features of equilibrium behavior in a specialized environment. Further, we compute the best and worst equilibrium value numerically for a range of parameters, including moderate discount factors.

B.1 Asymptotic Efficiency

For each behavioral type $\hat{\pi} \in \Pi_B$, define

$$\begin{aligned} \overline{v}^{\hat{\pi}} = \max \quad & \sum_{\theta \in \Theta} \sum_{m \in M} \mu_0(\theta) \hat{\pi}(m|\theta) v(\theta, a_m) \\ \text{s.t.} \quad & a_m \in A^*(\mu_m^{\hat{\pi}}) \text{ for all } m \in M, \end{aligned}$$

to be the rational sender's superior payoff, if he follows strategy $\hat{\pi}$ and the receiver knows this. The inferior payoff, $\underline{v}^{\hat{\pi}}$, is defined similarly with min instead of max. Note that these two values can differ only if the receiver has multiple optimal actions for some m.

Proposition 3. *For each $\beta \in ri(\Delta(\hat{\Pi}))$ and $\epsilon > 0$, there exists $\underline{\delta} < 1$ such that for all $\delta \geq \underline{\delta}$,*

$$\max_{\hat{\pi} \in \Pi_B} \underline{v}^{\hat{\pi}} - \epsilon \leq v \leq v^* + \epsilon \quad \textit{for all } v \in V(\beta; \delta).$$

In all PBE, the rational sender receives at least the inferior payoff associated with the best type he could imitate, $\max_{\hat{\pi}} \underline{v}^{\hat{\pi}}$, and at most his information-design value v^*, both approximately. This result follows from Theorem 1 in Gossner (2011) and Fudenberg and Levine (1992). Their results hold for a rather general extensive form stage game, and hence the choice of state each period in our model can be accommodated as a random move of nature (observed by the sender but not the receiver) in that extensive form.

What does this say about the ability of reputational enforcement to take the place of legal commitment? Information design relies on legal commitment but also assumes that the receiver plays the sender's favorite action when he has multiple best responses to a message. The standard argument in support of the latter is that for generic information design problems, there will be information structures (or signals) near the optimal one for which the receiver is never indifferent and that give the sender a payoff near v^*. Hence, without any tie-breaking assumptions, commitment by the sender can guarantee himself virtually his information-design value v^*. Consider such an information design problem, and an information structure that induces only unique best responses from the receiver and gives the sender $v^* - \epsilon/2$ for small $\epsilon > 0$. In any reputational game including that information structure, as the sender becomes patient, he is guaranteed virtually his information-design value in all PBE: for any $\beta > 0$, there exists $\underline{\delta} < 1$ such that

$$V(\beta; \delta) \subset [v^* - \epsilon, v^* + \epsilon] \quad \text{for all } \delta \geq \underline{\delta}.$$

B.2 Discussion

The welfare consequences of problems with limited observability have long been studied in the repeated game literature. Some systematic themes emerge that cast dramatic light on the result in the preceding paragraph. In Radner, Myerson and Maskin (1986), two players in a symmetric repeated partnership do not have enough information to punish someone who did not work hard without punishing the other partner as well: low output does not point to anyone in particular. The simultaneous moral hazard arising for both players, and the lack of outcomes that distinguish between the respective efforts of the two, mean that surplus must be thrown away in case of a bad outcome (by following a continuation equilibrium that is not ideal for either player). Because this inefficiency is generated by incentive constraints in every period, the inefficiency is "capitalized in the equilibrium value set", and average payoffs are uniformly bounded away from efficiency no matter how patient the players are.[24]

Contrast this with the repeated principal-agent problem in Radner (1985). Again, there is a moral hazard problem resulting from the agents action being unobservable, that must be faced every period. However, as the players become very patient, their interaction becomes highly efficient, unlike what happened with the symmetric partnership. This is made possible by surplus from their interaction being passed back and forth between the two players. When the agent gets unlucky and output is low, surplus is not destroyed, but rather passed to the principal. The principal pays the agent less than usual, but does not burn the money or donate it to charity: he pockets the difference. When instead there is a good realization of the random output, the agent is paid a lot; surplus is passed from the principal to the agent. So there is no destruction of surplus per period, as there was in the symmetric partnership. If the agent is risk averse, there is still the problem that he is not being fully insured by the risk neutral principal. This causes inefficiency for any fixed discount factor, but as the players become very patient, the consequences of an output failure today can be spread over many future periods, and asymptotically, equilibria are efficient. Fudenberg, Levine and Maskin (1994) prove in considerable generality that even in repeated games having more than one player with moral hazard problems, if there is enough information to discriminate between the players' behavior, again surplus can be passed back and forth, and very patient players can achieve highly efficient outcomes.

With these considerations in mind, think again about the NOC repeated game of Section 2. Because the long run player may not have myopic incentives to randomize according to the probabilities the equilibrium requires, and others cannot check this aspect of her behavior, there is a moral hazard problem each period. Surplus is not being passed back and forth: when the continuation equilibrium lowers the sender's payoff by moving to the babbling equilibrium for some time, this is bad for the receivers too. Unsurprisingly, the equilibrium has discounted average value for the sender that

[24]The same happens in more complicated settings with symmetric moral hazard (see Abreu, Milgrom and Pearce (1991), Kandori and Obara (2006), and Sannikov and Skrzypacz (2007)).

is lower than the KG legal commitment value, and the severity of this inefficiency is not relieved by any degree of patience. That is captured in Proposition A.2.

Finally, turn attention to the reputational model. The receiver still cannot see if the sender is generating the messages with the probabilities specified in equilibrium. The moral hazard issue remains, and must be faced each period. One would think that any equilibrium must display inefficiency in average discounted payoff that does not go away as the sender becomes sufficiently patient. And yet Proposition 3 says the opposite. *How can the cost of the moral hazard problem have been dismissed by the mere addition of low probability behavioral types?*

The answer lies in the fact that the reputation game is no longer a strictly repeated game. It is a dynamic game, indexed after any history of play by the sender's reputation. In effect, it has become an investment problem. The sender's inventory is his reputation for being behavioral. The continuation game he moves to depends on whether he consumes some of that inventory today or builds it up. Any time he sends a message that is myopically unattractive to him, he is building up that inventory, investing in his reputation. Randomization by the sender still requires indifference, and hence continuation values that compensate today's imbalance. But while these continuation values came from a fixed value set in the repeated game, implying welfare destruction for the sender, they come from value sets indexed by future reputations in the dynamic game. As the sender invests or disinvents period by period, he is effectively passing utility back and forth between his current and future selves.