REPUTATION AND INFORMATION DESIGN

Laurent Mathevet
European University Institute

David Pearce
New York University

Ennio Stacchetti
New York University

February 26, 2024

Abstract. Can reputation replace legal commitment for an institution making periodic announcements? Near the limiting case of ideal patience, results of Fudenberg and Levine (1992) imply a positive answer in *value* terms. However, because little is known about equilibrium *behavior* in dynamic reputational models, the classical dynamic foundation for commitment in Bayesian persuasion is incomplete. Computational and analytic approaches are combined here to characterize equilibrium behavior in a dynamic reputational cheap talk model. Behavior depends upon which of three reputational regions pertains after a history of play. These characterizations hold even far from the patient limit. But combined with a novel method of calculating average discounted values, they allow us to show behavioral convergence toward the static Bayesian persuasion solution.

†Special thanks are due to Vishal Ashvinkumar and Rumen Kostadinov for exceptional research assistance, and to Benjamin Bernard for his role as our discussant. We are also grateful to Dilip Abreu, Ben Brooks, Jeff Ely, 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 University of Arizona, the 2017 Latin American Meeting of the Econometric Society, the 2017 Becker Friedman Institute Economic Theory Conference, the 2020 North American Meeting of the Econometric Society, and the 2023 Northwestern conference honoring the memory of Nancy Schwartz.

Introduction

The publication of Kamenica and Gentzkow (2011) (KG, hereafter) saw an explosion of interest in Bayesian persuasion problems. In this static model, a sender commits to sending costless signals to a receiver as a function of information the sender will learn. In this way the sender may influence the receiver in her choice of action. The commitment assumption differentiates this "Bayesian persuasion" exercise from the cheap talk literature (Crawford and Sobel (1982)), allowing the sender to achieve a higher optimal value and theorists to provide powerful characterizations. A drawback is that commitment can be hard to justify in many situations, because of inherent interim opportunities to deviate.

Rayo and Segal (2010) proposed to employ Fudenberg and Levine (1989, 1992, FL hereafter) to justify the static payoff of a sender with ex-ante commitment as the equilibrium payoff of a long-run sender facing a sequence of short-run receivers in a dynamic reputational cheap talk game. Consider, for example, a car dealership selling a new car to a different buyer every period. In this dynamic setting, the reputational cost of dishonesty could potentially harm future sales and mitigate the short-term incentive to misrepresent the quality of a car. FL's results corroborate this notion, implying that a patient sender can secure an average discounted payoff virtually as high as his Bayesian persuasion payoff in any equilibrium. These reputational considerations are part of a broader effort to understand how repetition can substitute for the commitment assumption in Bayesian persuasion (see literature review).

Although the dynamic reputational model offers a classical foundation for Bayesian persuasion *payoffs*, the role of Bayesian persuasion *behavior* within this dynamic reputational narrative remains unclear. Indeed, FL obtain their results by using an imitation strategy that cannot be an equilibrium strategy,[1] and their convergence in value does not a priori imply convergence in behavior, as Bayesian persuasion payoffs could be obtained as average discounted payoffs without players ever adhering closely to Bayesian persuasion behavior. In fact, little is known about equilibrium behavior in dynamic reputational models. What is known, though, is that players will generally randomize, which is precisely why the standard repeated game often fails to reach the commitment payoffs in equilibrium![2] Thus, we find ourselves in the unsatisfying and rather puzzling situation where the classical reputational model lays the foundation for Bayesian persuasion payoffs, but we do not know *how*.

[1]If imitating the reputational type in every period were an equilibrium, there would be no way of distinguishing the rational type's behavior from the reputational type's, hence reputation would be constant and the sender always trusted. In this context, the rational type would lie with certainty when needed, contradicting said equilibrium behavior.

[2]In the standard repeated game (without reputation) between a long-run sender and a sequence of short-run receivers, randomization by the sender requires the continuation values to be adjusted to maintain indifference, which can induce a per period cost that does not arise in the static problem with commitment. 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) and Best and Quigley (2022)).

In a dynamic, reputational, cheap talk game with binary actions and states,[3] we characterize the equilibrium behavior that delivers the largest and the lowest equilibrium values, respectively, to the sender.[4] This analysis describes the dynamics of reputational management by the sender and the evolution of trust by the receiver. It also explains the mechanism by which both can approach their Bayesian persuasion payoffs for high discount factors, despite the sender's need to randomize between honest and dishonest reports. In these equilibria, the reputation space is divided into three regions, characterized by qualitatively uniform behavior within, but providing incentives in different ways across regions. In the very high reputation region, the sender capitalizes on his established reputation by lying with certainty, while the receiver trusts any recommendation. Outside that region, the sender faces the choice of investing or disinvesting in reputation, resolved stochastically. On the opposite end, in the low reputation region, the sender's probability of investing decreases with reputation, following a simple monotonic formula.

The intermediate reputation region features a notable phenomenon, as the sender's incentives are met by using *optimal* continuation values. This allows indifference without surplus destruction, and evokes the efficiency debate from repeated games with limited observability. Efficiency hinges on the ability to transfer surplus without waste between agents across periods. In Radner's (1985) repeated principal-agent problem, for instance, the principal compensates the agent less than usual after a low output, effectively transferring surplus from the agent to the principal. By contrast, in the symmetric repeated partnership model of Radner, Myerson and Maskin (1986), low output does not identify which partner should be disciplined, and hence surplus must be thrown away to deter shirking. As the sender invests or disinvests in his reputation inventory in our model, he is effectively passing utility back and forth between his current and future selves.

Our analysis is greatly simplified when the best equilibrium value for the sender is increasing in his reputation. If the upper and lower boundaries of the value set are too close together, characterizations of the upper boundary are potentially highly complex, because the lower boundary can "get in the way" of punishing dishonest behavior most effectively. By adapting the APS algorithm to the current problem, we provide numerical evidence, for a rich set of parameters, that the lower boundary is nowhere close to binding. This allows us to offer an analytical proof of the monotonicity of the upper boundary.[5]

Our computational results also establish that the rate at which the sender exploits his reputation (that is, reports dishonestly) is everywhere (not just in the low region discussed above) increasing. We can then prove a *behavioral convergence result* that

[3]The model captures a canonical yet stylized class of dynamic cheap talk games, such as a seller periodically selling goods of uncertain quality, an infectious disease expert advocating for healthy practices, or KG's celebrated prosecuting attorney model.

[4]This characterization holds for discount factors away from one.

[5]In previous versions of the paper, we had not obtained this proof.

clarifies the role of Bayesian persuasion behavior within the dynamic reputational narrative. For any behavioral type, $\epsilon > 0$ 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 Bayesian persuasion behavior. As the behavioral type asymptotically approaches the Bayesian persuasion behavior, this result implies that players actually mimic (near-) Bayesian persuasion behavior for so long, in discounted terms, that their average discounted payoffs approach Bayesian persuasion payoffs. This behavioral convergence result holds for the sender-preferred equilibrium, and approximately for values close to the upper frontier.

In the course of establishing behavioral convergence, we develop a novel technique called *stationary promising keeping*. When evaluating a player's average discounted payoff in a dynamic game, one can equivalently examine the payoff associated with any of the actions in the support of his current-period mixed strategy. The trouble is that this yields a weighted average of today's immediate payoff and a continuation value from tomorrow onward; often little is known about the latter. Instead, we focus on what happens if the sender behaved as though he had a target reputation, always lying if his reputation is above the target and telling the truth otherwise. Even though this is not equilibrium behavior, it is optimal behavior and can be used to establish a close relationship between the sender's payoff and his equilibrium rate of exploitation. In this paper, it allows us to prove behavioral convergence and to get a quick understanding of many asymptotic phenomena. We hope that stationary promising keeping may prove useful in other dynamic settings.

Our approach illustrates the complementary roles of theory and computation advocated by Judd (1997). The computational analysis provides an overview of properties away from the patient limit and the basis for behavioral convergence, thus clearing the path for a study of equilibrium behavior in reputational models. Numerical methods have been crucial in areas where dynamic and stochastic elements challenge the derivation of analytical solutions, such as dynamic stochastic general equilibrium (DSGE). See, for example, Gourio (2012, 2013), Andreasen (2012), Isoré and Szczerbowicz (2013, 2015), and Petrosky-Nadeau, Zhang, and Kuehn (2015). Applications of these methods extend to dynamic public finance (Golosov et al., 2007) and dynamic asset pricing (Borovicka and Stachurski, 2021). For an extensive overview of computational applications in economics, including dynamic programming with multiple state variables, non-linear DSGE models with shocks, and problems with binding constraints, readers are directed to resources on high-performance computing in economics, notably Aldrich, Fernández-Villaverde, Gallant, and Rubio-Ramirez (2010), Fernández-Villaverde and Levintal (2018), and Fernández-Villaverde and Valencia (2018).

The paper is organized as follows. Section 2 lays out the model. Section 3 presents the value convergence result. Section 4 introduces the essential incentive structure of perfect Bayesian equilibria of the model, and states the generalizations of self generation and the APS algorithm (Abreu, Pearce and Stacchetti (1990) that support our analysis. Using a robust computational result for the parameter ranges we consider, it also proves a monotonicity result for the upper frontier of the value correspondence: the best payoff increases with reputation. Section 5 then shows how behavior in the sender-preferred equilibrium varies across three regions of reputation space. Section 6 presents the behavioral convergence result, featuring a new technique we call stationary promise-keeping. The Appendix gives details of the computational method and contains the proofs. The computer codes that generate our numerical results and figures can be found at https://tinyurl.com/yc6eantw.

2. The Model

A long-lived sender communicates information to a sequence of short-lived receivers 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.[6]

At the beginning of period t, state θ_t is drawn from $\Theta = \{\ell, h\}$ according to distribution $\mu_0 \in \Delta\Theta$. When there is no confusion, denote $\mu_0 = \mu_0(h)$. The sender observes the realized state and then sends a message $m \in M = \{L, H\}$ 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). The behavioral type follows the same strategy $\pi_B : \Theta \to \Delta(M)$ in every period.

The receiver begins period t with a belief $\beta_t \in [0, 1]$ that the sender is s_B, which represents the sender's reputation, and with belief μ_0 about the state. Upon receiving message m_t, the receiver updates her belief about the state, resulting in posterior belief $\mathbb{P}[\theta|m_t, \beta_t, \pi_t]$. This updated belief also depends on two factors: first the sender's reputation β_t, which plays an important role in assessing the message's credibility, and secondly, the sender's strategy, π_t, which is known in equilibrium. Then, 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 an action to

$$\max_{a \in A} \sum_{\theta \in \Theta} \mathbb{P}[\theta|m_t, \beta_t, \pi_t] u(\theta, a).$$

The receiver always adopts a myopic best response to her beliefs in each period. This could be because she is inherently short-lived, as a buyer making a purchase decision

[6] 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.

before being replaced in the subsequent period. Alternatively, 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 sender, her individual actions are not detectable and hence she plays myopically.

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

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

The sender maximizes her average discounted expected payoff, where $\delta \in (0, 1)$ is the discount factor.

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

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

Let $V(\beta)$ be the set of PBE payoffs for the rational sender given the current receiver's belief β. Sometimes we write $V(\beta, \delta)$ to make explicit the dependence on the discount rate δ. 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 $[0, 1]$.

Throughout the paper, we assume:

Assumption: $\mu_0 < 1/2$.

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

Assumption (Trusted behavioral type): $\pi_B(H|h) = 1$ and $\pi_B := \pi_B(H|\ell) < \frac{\mu_0}{1-\mu_0}$.

The last assumption requires that the behavioral type must always tell the truth in the high state. It also restricts his probability of lying in the low state. Thus, a receiver who knows that the sender is s_B would rationally follow his advice, because it is the most likely state.

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 car dealership selling a new car to a different buyer every period. The seller only gets paid when he sells a car, while a buyer prefers to acquire a vehicle of high quality rather than one of inferior quality.

3. Asymptotic Efficiency

This section presents the value convergence result of FL within the specific context of our dynamic sender-receiver game.

Based on the trusted type assumption, if the sender were known to use strategy π_B in a given period, then his ex-ante expected value in that period would be

$$v^{\pi_B} \;=\; \mu_0 + (1-\mu_0)\pi_B.$$

In static Bayesian persuasion, a sender who commits to recommending H with certainty in state h and with probability

$$\boxed{\pi^* := \frac{\mu_0}{1-\mu_0}}$$

in state ℓ achieves an ex-ante expected value of

$$v^* \;=\; \mu_0 + (1-\mu_0)\pi^* \;=\; 2\mu_0.$$

The optimality of v^* for $\mu_0 < 1/2$ can be demonstrated by conventional techniques, such as Kamenica and Gentzkow (2011).[7]

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 $\delta \geq \underline{\delta}$.

This result is an immediate corollary of FL. In all PBE, the rational sender receives at least v^{π_B} and at most his Bayesian persuasion payoff, both approximately. It is easy to see that $\lim_{\pi_B \uparrow \pi^*} v^{\pi_B} = v^*$, so that a patient sender in the dynamic reputational game is virtually guaranteed, in all PBE, the same value as he would receive in the static Bayesian persuasion model.

This proposition is obtained by applying the imitation strategy of FL. In any PBE, the sender could deviate from his actual equilibrium strategy, and behave in the same way as the behavioral type by adopting strategy π_B in every period forever. This deviation would yield a patient sender a payoff of approximately v^{π_B}. Since no unilateral deviation can be strictly profitable in equilibrium, $V(\beta,\delta)$ must be at least as large. Although this imitation strategy is useful in establishing equilibrium payoff properties, it cannot be an equilibrium strategy (refer to Footnote 1 for further details).

In trying to understand what equilibrium behavior can support the conclusion of Proposition 1, one encounters a conundrum. The sender must sometimes recommend H in state ℓ to approach v^*, yet this recommendation must be probabilistic to maintain credibility. This requires an indifference condition in average discounted payoffs

[7] The case $\mu_0 \geq 1/2$ is trivial, because $v^* = 1$.

between recommending H and L—a condition that paradoxically impedes achieving v^* in the standard repeated games without reputational concerns (refer to Footnote 2 for further details)! Proposition 1 therefore introduces further questions about what self-enforcing mechanisms underlie convergence in value. In any case, since v^* could emerge from average discounted payoffs without players ever adhering closely to Bayesian persuasion behavior, the proposition alone offers limited insight into equilibrium behavior.

In Section 6, we present a stronger value convergence result: holding π_B fixed, the value of the sender-preferred equilibrium converges pointwise, as $\delta \to 1$, to a particular value, independent of β. Once again, this alone will *not* imply convergence of behavior across the reputation space. One could be forgiven for conjecturing that behavior converges to π_B. But it does not. Section 6 proves (Proposition 4) that it converges to the Bayesian persuasion behavior π^*.

4. Incentives and Computation

The tools of strategic dynamic programming, in particular self-generation, enable a characterization of equilibrium behavior as a function of the sender's reputation, even for discount factors away from the limit. At the heart of this methodology is our application of the APS algorithm (Abreu, Pearce, and Stacchetti, 1990) to reputational models, developed in Appendix A. The modified algorithm allows for detailed computational analyses to see what payoffs and behavior look like across diverse combinations of behavioral types and discount factors. This approach also offers strong evidence that, in the sender-preferred equilibrium and within the parameter ranges under consideration, the lower boundary of the value correspondence never hinders punishing the sender as efficiently as possible, when he tells a lie. This fact lets us give an analytic proof that the upper boundary of the equilibrium value correspondence is increasing with reputation, facilitating much of the analysis in the paper.

4.1. Reputational Incentives

After any history of play, a rational sender who observes today's state must decide whether to report truthfully or to lie. The optimal decision takes account of today's myopic payoff, how his possible reports will affect his reputation, and how those will affect his continuation payoffs from tomorrow onward. The relevant information is summarized in the triple:

1. Current reputation β, current strategies ($\pi(m|\theta)$ for the rational sender and $\alpha = (\alpha_H, \alpha_L)$ for the receiver, where $\alpha_m(a)$ describes the probability of playing action a after receiving message m), current average discounted value.
2. Continuation value and reputation tomorrow after a truthful message today
3. Continuation value and reputation tomorrow after a dishonest message today

Given today's message m and state θ, tomorrow's reputation $\beta_{m,\theta}$ is derived in (1) from today's reputation β. Today's and tomorrow's values must provide incentives for players to take the equilibrium actions. Formally, given θ, continuation values $w_{L,\theta} \in V(\beta_{L,\theta})$ and $w_{H,\theta} \in V(\beta_{H,\theta})$, one following honesty and the other a lie, and the receiver's strategy α, the rational sender will send an optimal message. That is,

$$\pi(m|\theta) > 0 \;\Rightarrow\; m \in \operatorname*{argmax}_{m'} \, (1-\delta) \sum_{a} \alpha_{m'}(a) v(a) + \delta w_{m',\theta} \tag{2}$$

In equilibrium, the receiver will also choose an optimal action for each message:

$$\alpha_m(a) > 0 \;\Rightarrow\; a \in \operatorname*{argmax}_{a'} \sum_{\theta} \mathbb{P}[\theta|m, \pi, \beta] u(\theta, a'). \tag{3}$$

These conditions form the definition of admissibility.

Definition. The tuple (π, α, w) is admissible at β if for all m, a and θ, $w_{m,\theta} \in V(\beta_{m,\theta})$, $\pi(m|\theta)$ satisfies (2) and $\alpha_m(a)$ satisfies (3).

We will prove that in the relevant equilibria, the rational sender's advice always remains trusted, so that $\alpha_H(H) = \alpha_L(L) = 1$, and that for most reputations β, the sender will randomize between honesty and lying in state ℓ. The latter requires indifference at $\theta = \ell$, between a payoff of 1 today with a continuation value $w_{H,\ell} \in V(\beta_{H,\ell})$ and a payoff of 0 today with a continuation value $w_{L,\ell} \in V(\beta_{L,\ell})$. That is,

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

Both sides of the equation, respectively referred to as left and right promise keeping, provide the same evaluation of the average discounted payoff in state ℓ. We rearrange the terms and express indifference as

$$w_{L,\ell} - w_{H,\ell} = \frac{1-\delta}{\delta}. \tag{4}$$

4.2. Monotone Optimal Value: Reputation as an Asset

At each reputation value, it is of particular interest to see how to optimize the sender's equilibrium payoff. Define

$$\underline{V}(\beta) = \min V(\beta) \quad \text{and} \quad \overline{V}(\beta) = \max V(\beta) \quad \text{for } \beta \in [0,1]$$

to be the upper and the lower boundaries of the equilibrium value correspondence. The optimization exercise underlying $\overline{V}$ and $\underline{V}$ motivates the subject of optimal admissible tuples.

Definition: A tuple (π, α, w) is optimal at β if it is admissible at β and

$$\sum_\theta \mu_0(\theta) \sum_m \pi(m|\theta) \left[(1-\delta) \sum_a \alpha_m(a) v(a) + \delta w_{m,\theta} \right] = \overline{V}(\beta).$$

To ensure the sender is incentivized to report truthfully, particularly after observing $\theta = \ell$, it may be necessary to implement punitive measures for dishonesty that result in a continuation value $w_{H,\ell}$ strictly below $\overline{V}(\beta_{H,\ell})$. This, in fact, might not even suffice to deter dishonest behavior adequately. In such situations, the temptation to cheat would need to be softened by reducing the receiver's likelihood of following advice below 1. However, this introduces significant inefficiencies, as it inadvertently punishes message H even when it is truthful!

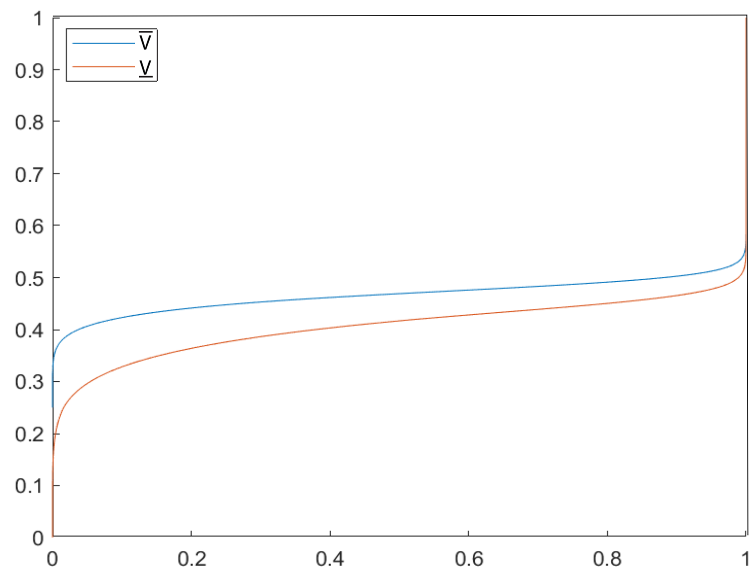

Figure 1: Value Correspondence for $\mu_0 = 0.25$, $\pi_B = 0.2$ and $\delta = 0.98$.

We investigate conditions under which these drastic measures can be avoided, focusing on the adjustment of continuation values as the only means of providing the desired incentives. Suppose the probability of lying in state ℓ, at current β, is increased to the point where the receiver is indifferent to following advice, because $\mathbb{P}[h|H, \beta, \pi] = 1/2$. It can be shown that this will generate updated reputation $\lambda_0\beta$ or $\lambda_1\beta$, respectively, after a dishonest or honest statement, where

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

The "gap condition" below guarantees that extreme measures, which would require the receiver to probabilistically ignore the sender's recommendations, are unnecessary.

Importantly, when the gap condition is satisfied, it not only obviates the need for overly punitive measures, but it also enables us to prove that the sender-preferred equilibrium value is weakly increasing in reputation, so that reputation is an asset for the sender.

Gap Condition: $\overline{V}(\lambda_1\beta,\delta) - \underline{V}(\lambda_0\beta,\delta) \geq (1-\delta)/\delta$ for all $\beta \in [0, 1/\lambda_1]$.

Using the modified APS algorithm from Appendix A, we compute the value correspondence and associated behavior for all combinations of parameters on the following grid:

$$\text{Grid} = \Big\{(\mu_0,\delta,\pi_B): \quad \mu_0 \in \{.1,.15,\ldots,.4\}, \\ \delta \in \{.925,\ldots,0.995,0.999\}, \pi_B \in \{0.02,0.04,\ldots,\bar{\pi}_B(\mu_0)\}\Big\}$$

where $\bar{\pi}_B(\mu_0)$ is the largest two-decimal place number below $\mu_0/(1-\mu_0)$. The algorithm lets us look at the value correspondence for any desired combination of parameters, as illustrated in Figure 1. For all parameter combinations within the grid, the algorithm also determines that the gap condition holds, and notably, it does so while leaving untapped at least 45% of the punitive capacity. In other words, the lower boundary is not even close to binding.

Numerical Property 1. [Gap Condition] For all (μ_0,δ,π_B) on the grid and $\beta \in [0,1]$,

$$\overline{V}(\lambda_1\beta,\delta) - \underline{V}(\lambda_0\beta,\delta) \;\geq\; (1-\delta/\delta \,+\, 0.45\,\big(\overline{V}(\lambda_0\beta,\delta) - \underline{V}(\lambda_0\beta,\delta)\big).$$

We maintain the Gap Condition as an assumption throughout the paper.

Proposition 2. [Monotonicity] If V satisfies the gap assumption, then $\overline{V}(\beta)$ is weakly increasing in β.

This result establishes a consistent relationship between maximal value and reputation, something that plays a central role in replacing legal commitment with reputational mechanisms.

What form PBEs take ultimately must depend on the parameters of the model. For example, if δ is reduced more and more, all intertemporal incentives will eventually collapse. Before that, presumably drastic forms of punishment will have to be used to supplement more efficient ones. When we present computational results here and in Section 5, we are not suggesting they will hold for all conceivable parameter values, but rather, for all those on our fairly rich grid.

5. Equilibrium Behavior Across Reputation Space

This Section reveals 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.

Before focusing on particular regions in the interior of reputation space, we discuss some basic properties of optimal tuples, and take care of two cases of reputational extremes.

5.1. Properties of Optimal Tuples

We characterize equilibrium behavior with initial value on the upper boundary by showing that the interior of reputation space can be organized into three regions. But first, consider two reputational extremes: $\beta = 0$ and $\beta = 1$.

When $\beta = 0$, it is common knowledge that the sender is not reputational, but rational. Here the dynamic game reduces to a standard game with one long run player interacting with a short run player each period. There is a babbling equilibrium in which the sender's messages are uninformative, and he is never trusted; its payoff to the seller is 0. Under our assumption on the discount factor, there is also a truthtelling equilibrium, in which the sender is trusted, but if he lies even once, he is never trusted again. This has value μ_0. A value recursion shows that there are no equilibria with higher values than this, no matter how high δ is (Lemma 4). Recall that with the commitment power of Bayesian persuasion, the sender could attain expected payoff of $v^* = 2\mu_0$. This is a dramatic expression of the possible inefficiency of supporting randomization in repeated games (see Fudenberg, Kreps and Maskin, 1990).

If instead $\beta = 1$, the receiver is certain that the sender is behavioral. Since the behavioral sender lies infrequently enough to be trusted, any recommendation to play H is respected, so the sender, if in fact rational (contrary to the receiver's view), can achieve a payoff of 1 in every period by claiming the state is high, without ever losing reputation. So $V(1) = \{1\}$.

The next proposition summarizes the results of Lemmas 8 to 10 in the Appendix for optimal tuples (or more properly, their translations from general W correspondences with monotonic upper frontiers, to the equilibrium value correspondence V). We state the proposition formally and then give a relatively informal account of what it says about behavior.

Proposition 3. Let (π, α, w) be an optimal tuple at β. Then:

1. $\pi(H|h) = 1$ and $\alpha_L(L) = 1$.
2. $\pi(H|\ell) \leq \bar{\pi}(\beta)$ and $\alpha_H(H) = 1$.
3. Without loss of generality, $\pi_B \leq \pi(H|\ell)$.

First, in the high state, the sender is always truthful, as recommending L instead would result not only in a zero payoff today but also annihilate reputation. This choice must be suboptimal, in line with our monotonicity theorem. Since recommendation L is exclusively associated with state ℓ, it is rational to always trust it, hence $\alpha_L(L) = 1$. Truthtelling in the high state also allows one to simplify notation: for any optimal tuple, denote

$$\boxed{\pi(\beta) := \max\Big\{\pi(H|\ell) : (\pi, \alpha, w) \text{ is an optimal tuple at } \beta\Big\}.}$$

Second, the sender does not lie so much that the receiver is unwilling to follow his advice about the state. There can be tuples in which that happens, but they are not optimal: one can always design a better tuple where the sender lies less, and therefore has some influence on the receiver.

More than this, optimality requires, under the gap condition which we maintain throughout (and which is easily satisfied across our entire parameter grid), that the receiver always follows advice with probability 1. We return to this point while discussing Region 1 below.

The receiver is willing to follow advice H if the probability that it signals $\theta = h$ is at least $1/2$. Indifference on her part therefore requires that $\mathbb{P}[h|H, \beta, \pi] = 1/2$, which is the case if

$$\pi(\beta) \;=\; \bar{\pi}(\beta) := \frac{\pi^* - \pi_B \beta}{1 - \beta}.$$

In every optimal tuple, the receiver always follows advice with probability 1, either because she is indifferent (when $\pi(\beta) = \bar{\pi}(\beta)$) or it is her unique best response (when $\pi(\beta) < \bar{\pi}(\beta)$).

Without loss of generality we can restrict attention to optimal tuples in which the rational sender lies more, after observing the low state, than does the behavioral type, that is, $\pi(\beta) \geq \pi_B$. Otherwise, a lie would result in a higher payoff today and a higher reputation tomorrow. Because $\pi(\beta) \geq \pi_B$, a lie after observing ℓ moves the reputation to the left, whereas an honest message moves it to the right.

5.1.1. Region 3: Exploitation

At an extremely high reputation, the sender is no longer willing to invest in his reputation by reporting honestly. His continuation payoffs from lying or telling the truth are both so close to 1 that the difference between them cannot prevent him from grabbing the myopic reward from lying today. We say that such a reputation β is in Region 3.

More generally, if setting $\pi(\beta) = 1$ and using continuation values on the upper frontier leaves the sender at least preferring to lie after observing ℓ, β is in Region 3. In summary, Region 3, the simplest of our regions, is the interval $[\beta_{23}, 1]$ where

$$\beta_{23} := \min\Big\{\beta : \ (\pi, \alpha, w) \text{ is optimal at } \beta \text{ and } \pi(\beta) = 1\Big\}.$$

5.1.2. Region 2: Efficient Incentives

Below β_{23}, the sender must always be randomizing. If instead he were supposed to send a pure message m after observing ℓ, deviating to the other message m' would prove him behavioral, giving him a reputation of 1 and continuation value of 1 tomorrow. This would be irresistible, contradicting the fact that m was supposed to be his best response.

The interval $(0, \beta_{23})$ is divided into regions we call Region 1 and Region 2, distinguished by the nature of punishments for lying. If, in an optimal tuple at reputation β, the continuation values $w_{L,\ell}$ and $w_{H,\ell}$ are both on the upper frontier, we say β is in Region 2. When the sender lies in state ℓ, he loses reputation, but still receives the highest payoff available at his new lower reputation. In this sense, the punishments for lying are not wasteful: surplus is not thrown away by using continuation values below the upper frontier of V. Instead, as the sender sometimes lies and sometimes advises honestly, he is effectively passing surplus back and forth between his current and future selves, as he draws upon or adds to his stock of reputation. There is room for this mechanism in the current game, exactly because it is a dynamic game rather than a strictly repeated game (the class for which Fudenberg, Kreps and Maskin (1990) produced a uniform inefficiency result).

Think of designing an optimal tuple at β in Region 2. As you gradually increase $\pi(\beta)$, reputation after honesty increases and reputation after a lie decreases. Because $\overline{V}$ is increasing (monotonicity theorem), this growing difference in posterior reputations means a greater penalty in terms of continuation values tomorrow for lying today. Stop increasing when (4) holds exactly; this makes the sender who observes ℓ indifferent between lying and telling the truth, and hence willing to randomize as necessary for equilibrium.

What if, as you raise $\pi(\beta)$, it hits $\bar{\pi}(\beta)$ before the incentive constraint is satisfied? That is, raising $\pi(\beta)$ enough to satisfy (4) makes the receiver unwilling to follow advice. Then keeping both continuation values on the upper frontier of V is incompatible with satisfying incentives for both the sender and the receiver, and we say β is in Region 1, which we discuss below. Let

$$\underline{\beta} = \inf\big\{\beta : \ (\pi, \alpha, w) \text{ is an optimal tuple at } \beta \text{ and } \pi(\beta) < \min\{\bar{\pi}(\beta), 1\}\big\}$$

be the infimum of Region 2.

Except at the boundary of Region 2, $\pi(\beta) < \bar{\pi}(\beta)$, and the receiver strictly prefers to follow advice. So $\alpha_H(H) = 1$. Both behavioral and rational types of sender report truthfully when the state is h, and the receiver follows the advice.

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

5.1.3. Region 1: Inefficient Incentives

An optimal tuple delivers incentives to the sender as efficiently as possible. In Region 2, one can have advice always followed and the sender indifferent between honesty and lying in the low state, without using continuation values below the upper frontier of V. By definition, this is not possible in Region 1. Although $\pi(\beta) = \bar{\pi}(\beta)$ to maximize spread, the incentives created for honesty are inadequate. There are two ways to make lying less attractive to the sender: lowering the continuation value after a lie, or having the receiver follow a recommendation of H only probabilistically (thereby reducing the sender's immediate reward today from lying). Intuitively, the second method is the more wasteful: it is poorly targeted, punishing message H whether it is true or not. Lemma 13 proves that if $w_{L,\ell}$ can be dropped below the upper frontier to meet the incentives in (4), without violating the lower boundary of V and without setting $\alpha_H(H) < 1$, then this is the optimal configuration. Reducing $\alpha_H(H)$ is a last resort.

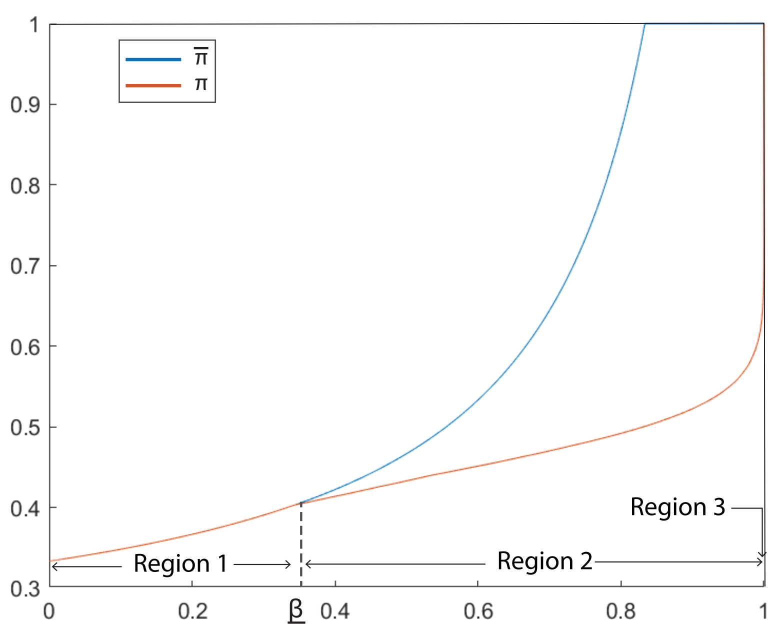

Figure 2: Sender's Equilibrium Strategy ($\mu_0 = 0.25$, $\pi_B = .2$ and $\delta = .98$)

Because of the gap condition from Subsection 4.2, which is satisfied effortlessly across our parameter grid (Numerical Property 1), it is never necessary to set $\alpha_H(H) < 1$ at any reputation β.

Nothing in the analytical results ensures that Regions 1 and 2 are intervals. Alternations between the two are conceivable, where the continuation value after a dishonest report might lie below the upper boundary. Interestingly, the regions have the simplest possible structure for parameter values on the grid:

Numerical Property 2. Regions 1, 2 and 3 are all intervals.

5.2. Lower Boundary Behavior

Start with the following numerical observation:

Numerical Property 3. For all (μ_0, δ, π_B) on the grid, $\underline{V}(\beta)$ is strictly increasing in β.

We make just a few remarks about the lower boundary. It too comprises three regions. 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 $\pi(H|\ell) = 1$ and the receiver follows his advice with certainty. 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 5.1. 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, $\alpha_H(H) = 0$, so that there would be no chance of H. But this requires $\pi(H|\ell) = \bar{\pi}(\beta)$ by a variant of Lemma 8. This may or may not be the best way of destroying value, as a larger $\pi(H|\ell)$ 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 $\alpha_H(H) = 0$. This being the case, to make the sender indifferent requires giving him equal continuation payoffs: $w_{H,\ell} = w_{L,\ell}$. Moreover, $\pi(H|\ell) = \bar{\pi}(\beta)$, 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), $\pi(H|\ell) < \bar{\pi}(\beta)$ and hence $\alpha_H(H) = 1$.

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 9, the receiver chooses H by Lemma 8, 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 the rational sender lies more frequently than does the behavioral type, there is an overall downward drift in reputation.[8] 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.

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.

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, but monotonicity 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 all (μ_0, π_B) on the grid, there exists $\underline{\delta} \in (0,1)$ such that for all $\delta \geq \underline{\delta}$, $\pi(\beta)$ is strictly increasing in reputation β.

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

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 15-20 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, $|\pi(\beta_t) - \pi^*| < \epsilon$ 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

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

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 $t = 0, \ldots, T$. 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 $\beta \in [\underline{\beta}_\epsilon, 1]$, $\bar{\pi}(\beta) \geq \pi^* + \bar{\epsilon} > \pi^* + \hat{\epsilon} \geq \pi(\beta)$.Since $1-\hat{\epsilon} > 1-\bar{\beta}_\epsilon$, the assumption made earlier is satisfied and indeed $\beta^L_t \leq \bar{\beta}_\epsilon$ for all $t = 0, \ldots, T$. 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 $t = 1, \ldots, T$, and thus $\pi^* < \pi(\beta_t) < \pi^* + \epsilon$ for all $t = 0, \ldots, T$. □

We make several remarks. First, convergence in behavior is not an immediate consequence of convergence in value (Proposition 1 and FL).[9] 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 18 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.[10]

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 $\pi(\cdot) \geq \pi_B$ by Proposition 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 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

[9]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.

[10]Away from the boundary, there are many admissible tuples that support the same value at a given reputation. Accordingly, Jeff Ely asked whether the behavioral convergence was a knife-edge phenomenon, in the sense that it holds only on the upper boundary of V. It is not as fragile as that: we can show that values close to the upper boundary entail behavior close to the Bayesian persuasion benchmark.

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 1. 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.[11]

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 17 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.

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

[11] 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.

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 $\Pi_B = \{\pi_B^1, \dots, \pi_B^K\}$ (each representing some probability of lying in state ℓ) are available[12] 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 $n = 1, \dots, K$.

In virtue of this assumption, all types in Π_B are trusted by the receiver.[13] 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:

Numerical Property 4. For $\delta = .95$, the rational sender does strictly better in equilibrium in the subgame following $\pi_B = 0.2$ than in the subgame following $\pi_B = 0.3$, $\overline{V}_{0.2}(\beta, \delta) > \overline{V}_{0.3}(\beta, \delta)$, for all reputations $\beta \in [0, 0.85]$.

[12] 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.

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

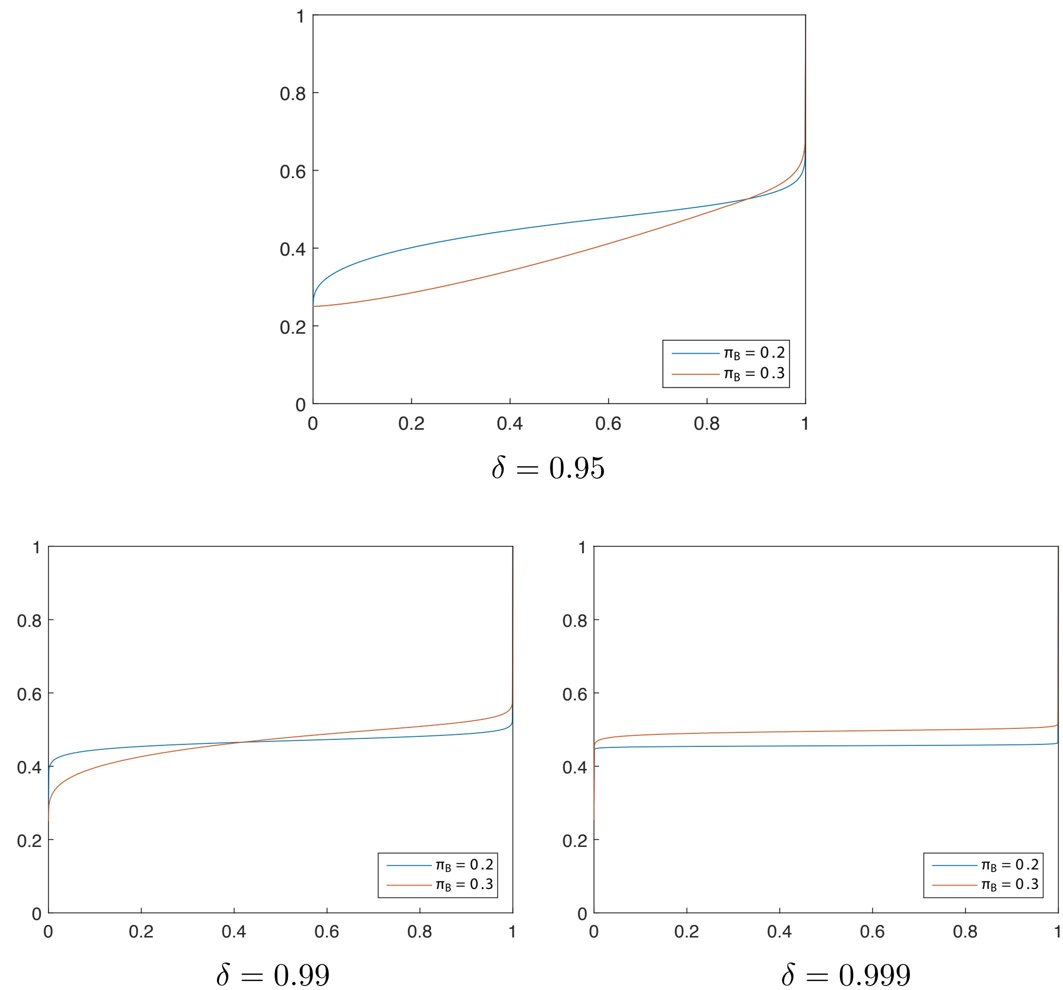

Figure 4: $\overline{V}$ for $\pi_B = 0.2$ and $\pi_B = 0.3$

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 .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 2 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 2 also shows the further consequence that Region 1 would then extend across almost the entire reputation space.

Lemma 2. 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, 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 (2023), 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 (2023) 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 Fudenberg, Kreps and Maskin (1990). 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. In subsection 8.1 below we discuss the relationships to our analysis in some detail. 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;[14] 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 (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.

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

He gives conditions under which the sender can attain the Bayesian persuasion payoff. Fudenberg, Gao and Pei (2022) give positive results for a novel problem in which a long run sender is partially informed of which actions will be available to him. Before the uncertainty is fully resolved, he announces his intended action; he can build a reputation for carrying out his announced intentions, when those are feasible.

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).

8.1. 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 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.

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.

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, we prove that, when equilibrium probabilities of lying increase with reputation, as they do everywhere on our parameter grid, 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.

Appendix

A. Reputational APS Framework

A.1. Self Generation

In equilibrium, continuation values are drawn from the equilibrium value correspondence V. But to employ the methods of strategic dynamic programming, it is useful to imagine those values being drawn from an abstract value correspondence W. Let $\mathcal{C}$ be the collection of correspondences $W : [0,1] \rightrightarrows \mathbb{R}_+$ with a compact graph such that $W(\beta)$ is a nonempty and compact interval for each $\beta \in [0,1]$. 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. For any $W \in \mathcal{C}$, let

$$\underline{W}(\beta) = \min W(\beta) \quad \text{and} \quad \overline{W}(\beta) = \max W(\beta) \quad \text{for } \beta \in [0,1].$$

Value correspondences W with a monotone upper boundary $\overline{W}$ play an important role in our analysis. Let $\mathcal{C}^+$ denote the set of correspondences $W \in \mathcal{C}$ such that $\overline{W}(\beta)$ is weakly increasing in β.

Fix $W \in \mathcal{C}^+$. It is possible that $\overline{W}$ is not continuous, but since W is upper hemicontinuous (as it has a compact graph), $\overline{W}$ is upper semicontinuous. As $\overline{W}$ is a weakly increasing function, $\overline{W}$ must also be right continuous. Define

$$\overline{W}_L(\beta) = \lim_{\beta' \uparrow \beta} \overline{W}(\beta').$$

By upper hemicontinuity, $\overline{W}_L(\beta) \leq \overline{W}(\beta)$ and $[\overline{W}_L(\beta), \overline{W}(\beta)] \in W(\beta)$.

Assume that in the current period the receiver believes the sender is behavioral with probability β (and rational with probability $1-\beta$), and expects the rational sender to play strategy $\pi : \Theta \to \Delta(M)$. After receiving message m, the receiver updates her beliefs about the state and chooses a myopic (possibly mixed) best reply $\alpha_m \in \Delta(A)$. Let $\mathrm{BR}^R(\beta, \pi) = \{\alpha = (\alpha_H, \alpha_L) \mid \alpha_m \text{ is a best reply to message } m,\ m \in M\}$. Since $A = \{H, L\}$ has only two actions, below we often write α_m to denote $\alpha_m(H)$.

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

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

Definition. *Given $W \in \mathcal{C}$, the tuple (π, α, w) is admissible for W at $\beta \in [0,1]$ 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 attainable payoffs given W:

$$B(W)(\beta) = \{E(\pi, \alpha, w) \mid (\pi, \alpha, w) \text{ is admissible for } W \text{ at } \beta\}, \quad \beta \in [0,1],$$

and define $\tilde{B}(W)$ as $\tilde{B}(W)(\beta) = \mathrm{co}(B(W(\beta)))$ for each $\beta \in [0,1]$. 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 5. 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.

Proposition 6. $V = \tilde{B}(V)$. Moreover, V is the largest fixed point of $\tilde{B} : \mathcal{C} \to \mathcal{C}$.

A.2. Algorithm and Existence

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 $W^\infty = V$.

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 3. If $W \in \mathcal{C}$ then $\tilde{B}(W) \in \mathcal{C}$.

The crucial step in the proof of Lemma 3 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.

Proof: Let $\underline{w} = \min\{\underline{W}(\beta) \mid \beta \in [0,1]\}$ and $\bar{w} = \max\{\overline{W}(\beta) \mid \beta \in [0,1]\}$ Clearly $E(\pi,\alpha,w) \in [(1-\delta)\underline{v} + \delta\underline{w}, (1-\delta)\overline{v} + \delta\overline{w}]$ for each $\beta \in [0,1]$ and each tuple (π,α,w) admissible for W at β. Therefore

$$B(W) \subseteq [0,1] \times [(1-\delta)\underline{v} + \delta\underline{w}, (1-\delta)\overline{v} + \delta\overline{w}]$$

is a bounded set. Let $\{(\beta^k, x^k)\} \subseteq B(W)$ be a sequence such that $\beta^k \to \beta$ and $x^k \to x$. We now show that $(\beta, x) \in B(W)$. For each $k \in \mathbb{N}$, there is a tuple (π^k, α^k, w^k) admissible for W at β^k such that $x^k = E(\pi^k, \alpha^k, w^k)$. Since $\Delta(M)^\Theta$, $\Delta(A)^M$ and $[\underline{w}, \overline{w}]^{M\times\Theta}$ 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 $w^\infty \in [\underline{w}, \overline{w}]^{M\times\Theta}$. One can check that $(\pi^\infty, \alpha^\infty, w^\infty)$ is admissible for W at β. Finally,

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

Therefore, $(\beta, z) \in B(W)$. This establishes that $B(W) \subseteq [0,1] \times \mathbb{R}$ is a closed set and hence that $B(W) : [0,1] \to [(1-\delta)\underline{v} + \delta\underline{w}, (1-\delta)\overline{v} + \delta\overline{w}]$ is an upper hemicontinuous correspondence. Since $\tilde{B}(W)(\beta) = \text{co}(B(W(\beta)))$ for each $\beta \in [0,1]$, $\tilde{B}(W)(\beta)$ is a compact convex set for each $\beta \in [0,1]$ and $\tilde{B}(W) : [0,1] \to [(1-\delta)\underline{v} + \delta\underline{w}, (1-\delta)\overline{v} + \delta\overline{w}]$ is also an upper hemicontinuous correspondence.

It remains to show that $B(W)(\beta) \neq \emptyset$ for each $\beta \in [0,1]$. Fix $\beta \in [0,1]$ 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":

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

Clearly, for each (m, θ), the function $\pi \to \beta_{m,\theta}(\beta, \pi)$ from $\Delta(M)^\Theta$ to $[0, 1]$ is continuous. Since $W \in \mathcal{C}$, the correspondence $\pi \to W(\beta_{m,\theta}(\beta, \pi))$ 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 $\Delta(M)^\Theta \times \Delta(A)^M \times [\underline{w}, \overline{w}]^{M\times\Theta}$ into itself has a fixed point (π, α, w). Therefore, $E(\pi, \alpha, w) \in B(W)(\beta)$ and $B(W)(\beta) \neq \emptyset$. □

Proposition 7. $W^\infty \in \mathcal{C}$ and $W^\infty = V$.

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

Proof: 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 $\beta \in [0, 1]$ 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 $\lambda^k \in [0, 1]$ and two tuples $(\pi^{1k}, \alpha^{1k}, w^{1k})$ and $(\pi^{2k}, \alpha^{2k}, w^{2k})$ admissible for W^k at β such that $w = \lambda^k E(\pi^{1k}, \alpha^{1k}, w^{1k}) + (1 - \lambda^k)E(\pi^{2k}, \alpha^{2k}, w^{2k})$. Again, without loss of generality we can assume that $\lambda^k \to \lambda^\infty$, $\pi^{jk} \to \pi^{j\infty}$, $\alpha^{jk} \to \alpha^{j\infty}$ and $w^{jk} \to w^{j\infty}$ for some $\pi^{j\infty} \in \Delta(M)^\Theta$, $\alpha^{j\infty} \in \Delta(A)^M$ and $w^{j\infty} \in [\underline{w}, \overline{w}]^{M\times\Theta}$. It is easy to check that $w^{j\infty}_{m,\theta} \in W^\infty(\beta_{m,\theta}(\beta, \pi^{j\infty}))$ 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,

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

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

By Proposition 5, $W^\infty \subseteq V$. 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 $V \subseteq W^\infty$. □

A.3. Monotone Optimal Value: Proofs

We begin with the degenerate case where reputation is 0: it is common knowledge that the sender is rational, not behavioral. The value set there includes 0, the sender's payoff in the babbling equilibrium, and also the truthtelling equilibrium, which relies on our assumption $\delta \geq 1/(1+\mu_0)$.

Lemma 4. $V(0) = [0, \mu_0]$ for all $\delta \geq 1/(1+\mu_0)$.

Proof: When $\beta = 0$ the game reduces to an infinitely repeated game with perfect information. In the infinitely repeated game there is a "babbling" equilibrium where the sender's messages are ignored and the receiver chooses action L in every period regardless of the sender's message. Given that the sender's messages are uninformative and that $\mu_0 < 1/2$, L is the unique optimal action for the receiver. This equilibrium has an expected value 0 (for the sender) and is the worst equilibrium for the sender. The repeated game also has an equilibrium where on the outcome path the sender always tells the truth and the receiver accepts all the sender's recommendations. That is, in every period t on the equilibrium path, the sender chooses the strategy π_t where $\pi_t(H|h) = 1$ and $\pi_t(L|\ell) = 1$, and the receiver chooses the action $a_t = H$ if $m_t = H$ and the action $a_t = L$ if $m_t = L$. If the sender ever lies, the players revert to the babbling equilibrium. This strategy has expected value μ_0 for the sender. Given that the receiver accepts all his recommendations, the sender is tempted to send the message $m_t = H$ when $\theta_t = \ell$. But the continuation value when he lies is $(1-\delta) + \delta 0$, while his continuation value from telling the truth is $\delta\mu_0$. This deviation is not profitable. A simple value recursion shows that this is an optimal equilibrium. Therefore $\{0, \mu_0\} \subset V(0)$ and $[0, \mu_0] = V(0)$. □

Define

$$\mathcal{C}_0 = \{W \in \mathcal{C} \mid W \supset V\} \quad \text{and} \quad \mathcal{C}_0^+ = \{W \in \mathcal{C}^+ \mid W \supset V\}.$$

In this section, we fix $W \in \mathcal{C}_0^+$ and let

$$X = \tilde{B}(W).$$

We now focus on the admissible tuples that support the optimal values $\bar{X}(\beta)$, $\beta \in [0,1]$.

Note that for any π and β, the posteriors

$$\beta_{H,\ell}(\beta,\pi) = \frac{\beta\pi_B}{\beta\pi_B + (1-\beta)\pi(H|\ell)} \qquad \text{and}$$

$$\beta_{L,\ell}(\beta,\pi) = \frac{\beta(1-\pi_B)}{\beta(1-\pi_B) + (1-\beta)(1-\pi(H|\ell))}$$

depend exclusively on $\pi(H|\ell)$. Hereafter we express these posteriors as $\beta_{H,\ell}(\beta,\pi(H|\ell))$ and $\beta_{L,\ell}(\beta,\pi(H|\ell))$ instead. Also, to simplify notation, we sometimes omit the arguments and simply write $\beta_{H,\ell}$ and $\beta_{L,\ell}$ instead of $\beta_{H,\ell}(\beta,\pi(H|\ell))$ and $\beta_{L,\ell}(\beta,\pi(H|\ell))$.

Lemma 5. $\overline{X}(0) = (1-\delta)\mu_0 + \delta\overline{W}(0) \geq \mu_0$.

Proof: If (π,α,w) is an admissible tuple for W at $\beta = 0$ and $\alpha_H(H) = 0$ (and hence $\alpha_L(H) = 0$ as well), then

$$E(\pi,\alpha,w) = \delta\big[\mu_0[\pi(H|h)w_{H,h} + \pi(L|h)w_{L,h}] + (1-\mu_0)[\pi(H|\ell)w_{H,\ell} + \pi(L|\ell)w_{L,\ell}]\big]$$

and $E(\pi,\alpha,w) \leq \delta\overline{W}(0)$. On the other hand, a tuple (π,α,w) with $\alpha_H(H) > 0$ is admissible only if the receiver's posterior that $\theta = h$ given the message H is greater or equal to $1/2$. That is,

$$\mathbb{P}[h|H] = \frac{\pi(H|h)\mu_0}{\pi(H|h)\mu_0 + \pi(H|\ell)(1-\mu_0)} \geq \frac{1}{2} \quad \text{or} \quad \pi(H|h)\mu_0 - \pi(H|\ell)(1-\mu_0) \geq 0.$$

Since $\mu_0 < 1/2$, this implies that $\pi(H|\ell) < 1$ and

$$\pi(L|h)\mu_0 - \pi(L|\ell)(1-\mu_0) = (2\mu_0 - 1) - [\pi(H|h)\mu_0 - \pi(H|\ell)(1-\mu_0)] < 0.$$

Thus $\mathbb{P}[h|L] < 1/2$. Therefore $\alpha_L(H) = 0$ and $E(\pi,\alpha,w) = \mu_0 v_h + (1-\mu_0)v_\ell$ where

$$v_\theta = \pi(H|\theta)(\alpha_H(1-\delta) + \delta w_{H,\theta}) + (1-\pi(H|\theta))\delta w_{L,\theta} \quad \text{for } \theta = h, \ell.$$

Note that when $\beta = 0$, $\beta_{m,\theta} = 0$ for all (m,θ). One can check that any tuple (π^o,α^o,w^o) where

$$\pi^o(H|h) = 1, \quad \pi^o(H|\ell) \leq \frac{\mu_0}{1-\mu_0}, \quad \alpha^o_H(H) = \alpha^o_L(L) = 1 \quad \text{and}$$
$$w^o = (w^o_{H,h}, w^o_{L,h}, w^o_{H,\ell}, w^o_{L,\ell}) = (\overline{W}(0), 0, 0, \overline{W}(0))$$

has value

$$E(\pi^o,\alpha^o,w^o) = (1-\delta)\mu_0 + \delta\overline{W}(0)$$

and is optimal for W at $\beta = 0$. □

Definition. *A tuple (π,α,w) is optimal for W at some $\beta \in [0,1]$ if it is admissible at β and $E(\pi,\alpha,w) = \overline{X}(\beta)$.*

We will show below that if (π,α,w) is an optimal tuple for W at some $\beta \in (0,1)$, then $\pi(H|h) = 1$. Let (π,α,w) be an admissible tuple for W at some $\beta \in (0,1)$ such that $\pi(H|h) = 1$. Then, by Bayes' rule, the receiver's posterior belief that $\theta = h$ given message H is

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

Let

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

and note that if $\pi(H|\ell) = \bar{\pi}(\beta)$, the receiver is indifferent about following the recommendation H because in this case $\mu(\pi, \beta) = 1/2$. When the sender deceives as often as can in state ℓ without compromising overall trust, that is, when $\pi(H|\ell) = \bar{\pi}(\beta)$, he induces updated reputations $\beta_{H,\ell} = \lambda_0\beta < \beta$ if the dishonest message is realized, and $\beta_{L,\ell} = \lambda_1\beta > \beta$ otherwise, where

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

Note that $\bar{\pi}(\beta_d) = 1$ when $\beta_d = 1/\lambda_1$.

In our numerical computations we always observe that V satisfies the gap condition for all the parameter specifications.

As specified in 4.3, the Gap Condition on V is maintained throughout. We show below that under this assumption, $\overline{V}(\beta)$ is weakly increasing in β.

Lemma 6. If $W \in \mathcal{C}_0$, then W satisfies the gap condition.

Proof: For any $\beta \in [0, 1/\lambda_1]$

$$\overline{W}(\lambda_1\beta) - \frac{1-\delta}{\delta} \geq \overline{V}(\lambda_1\beta) - \frac{1-\delta}{\delta} \geq \underline{V}(\lambda_0\beta) \geq \underline{W}(\lambda_0\beta). \qquad \square$$

We now construct an admissible tuple (π^o, α^o, w^o) for each $\beta \in [0,1]$. We will show later that these tuples are optimal. For any $\beta \in (0,1)$ and π such that $\pi(H|\ell) > \pi_B$, $\beta_{H,\ell}(\beta, \pi(H|\ell)) < \beta < \beta_{L,\ell}(\beta, \pi((H|\ell))$. Let

$$\hat{\pi}(\beta) = \sup\{\pi(H|\ell) \in [\pi_B, 1] \mid \delta[\overline{W}(\beta_{L,\ell}(\beta, \pi(H|\ell))) - \overline{W}_L(\beta_{H,\ell}(\beta, \pi(H|\ell)))] \leq 1-\delta\}.$$

Note that $\beta_{L,\ell}(\beta, 1) = 1$ and that $\beta_{H,\ell}(\beta, 1)$ is a continuously increasing function of β. Let

$$\beta_1 = \inf\{\beta \in [0,1] \mid \hat{\pi}(\beta) = 1\} \quad \text{and} \quad \beta_{23} = \min\{\beta_1, 1/\lambda_1\}.$$

Then

$$\overline{W}(1) - (1-\delta)/\delta \in [\overline{W}_L(\beta_{H,\ell}(\beta_1, 1)), \overline{W}(\beta_{H,\ell}(\beta_1, 1)]$$

and for all $\beta > \beta_1$, by the monotonicity of $\overline{W}(\beta)$, $\hat{\pi}(\beta) = 1$ and

$$\delta[\overline{W}(1) - \overline{W}_L(\beta_{H,\ell}(\beta_1, 1)] < 1 - \delta.$$

Moreover, if $\bar{\pi}(\beta) \geq 1$, following an H recommendation from the sender is optimal even if $\pi(H|\ell) = 1$.

For $\beta \in [0, \beta_{23})$, define the tuple (π^o, α^o, w^o) by

1. $\pi^o(H|h) = 1$ and $\pi^o(H|\ell) = \min\{\hat{\pi}(\beta), \bar{\pi}(\beta)\}$
2. $\alpha^o_H(H) = \alpha^o_L(L) = 1$
3. $w^o_{L,\ell} = \min\{\overline{W}(\beta_{L,\ell}(\beta, \pi^o)(H|\ell)), \overline{W}(\beta_{H,\ell}(\beta, \pi^o)(H|\ell)) + (1-\delta)/\delta\}$,
4. $w^o_{H,\ell} = w^o_{L,\ell} - (1-\delta)/\delta$,
5. $w^o_{L,h} = \underline{W}(0)$ (for convenience) and $w^o_{H,h} = \overline{W}(\beta)$.

When $\delta[\overline{W}(\beta_{L,\ell}(\beta, \pi^o(H|\ell))) - \overline{W}_L(\beta_{H,\ell}(\beta, \pi^o(H|\ell)))] > 1 - \delta$ (because $\pi^o(H|\ell) = \hat{\pi}(\beta) \leq \bar{\pi}(\beta)$ and the upper boundary of W has a vertical segment at $\beta_{H,\ell}(\beta, \pi^o(H|\ell))$ or at $\beta_{L,\ell}(\beta, \pi^o(H|\ell))$), the continuation values $w^o_{L,\ell}$ and $w^o_{L,\ell}$ are maximized while ensuring that

$$w^o_{m,\ell} \in [\overline{W}(\beta_{m,\ell}(\beta, \pi^o(H|\ell))) - \overline{W}_L(\beta_{m,\ell}(\beta, \pi^o(H|\ell)))] \quad m = L, H.$$

For $\beta \in [\beta_{23}, 1]$, define the tuple (π^o, α^o, w^o) by

1. $\pi^o(H|h) = \pi^o(H|\ell) = 1$
2. $\alpha^o_H(H) = \alpha^o_L(L) = 1$
3. $w^o_{L,\ell} = \overline{W}(1)$ and $w^o_{H,\ell} = \overline{W}(\beta_{H,\ell}(\beta, 1))$
4. $w^o_{L,h} = \underline{W}(0)$ (for convenience) and $w^o_{H,h} = \overline{W}(\beta)$

Lemma 7. For each $\beta \in [0, 1]$, the corresponding (π^o, α^o, w^o) is an admissible tuple for W at β.

Proof: Consider first the case when $\beta \in (0, \beta_{23}]$. Then the continuation values satisfy the incentive constraint:

$$(1-\delta) + \delta w^o_{H,\ell} = \delta w^o_{L,\ell}, \tag{IC}$$

the sender is indifferent between messages H and L when $\theta = \ell$ and thus any $\pi^o(H|\ell) \in [0, 1]$ is optimal for the sender. Also, since $w^o_{L,h} = \underline{W}(0) < \overline{W}(0) < \overline{W}(\beta) = w^o_{H,h}$, $(1-\delta) + \delta w^o_{H,h} > \delta w^o_{L,h}$ and the sender strictly prefers sending message H when $\theta = h$, so $\pi^o(H|h) = 1$ is optimal for the sender. As we remarked in Lemma XX, given that $\pi^o(H|h) = \pi_B(H|h) = 1$, the receiver strictly prefers to take action L after receiving message L and thus $\alpha^o_L(L) = 1$ is optimal for her. Finally, given that $\pi^o(H|\ell) \leq \bar{\pi}(\beta)$, the receiver (weakly) prefers to take action H after receiving message H and thus $\alpha^o_H(H) = 1$ is optimal for her. Note also that $w^o_{m,\theta} \in W(\beta_{m,\theta})$ for all (m, θ). In particular $\beta_{L,h} = 0$, and $w^o_{L,h} \in W(\beta_{L,h})$.

Now consider the case when $\beta \in (\beta_{23}, 1]$. Then $\hat{\pi}(\beta) = 1$. This implies that even when $\pi(H|\ell) = 1$, so that $\beta_{L,\ell}(\beta, 1) = 1$, $(1-\delta)+\delta\overline{W}(\beta_{H,\ell}(\beta, 1)) \geq \delta\overline{W}(1)$. Therefore, since $\alpha^o_H(H) = 1$, it is optimal for the sender to send the message H when $\theta = \ell$. Also it is strictly optimal for him to send message H when $\theta = h$. That $\pi^o(\beta) = 1$ also implies that $\bar{\pi}(\beta) \geq 1$, so that $\mu(\pi^o, \beta) \geq 1/2$ and $\alpha^o_H = 1$ is optimal for the receiver. As we saw before, since $\pi^o(H|h) = \pi_B(H|h) = 1$, $\alpha^o_L(L) = 1$ is optimal for the receiver. □

Corollary. For any $\beta \in (0, 1]$, $\overline{X}(\beta) \geq \overline{X}(0)$.

Proof: Fix $\beta \in (0, 1]$ and its corresponding (π^o, α^o, w^o). If $\beta \in (0, \beta_{23})$ then the (IC) is satisfied and the sender is indifferent between messages H and L when $\theta = \ell$. Therefore, the value of this tuple is

$$E(\pi^o, \alpha^o, w^o) = \mu_0[(1-\delta) + \delta w^o_{H,h}] + (1-\mu_0)\delta w^o_{L,\ell}. \qquad (RPK)$$

Recall that $w^o_{H,h} = \overline{W}(\beta)$ and that $\beta_{L,\ell} > \beta$ so $w^o_{L,\ell} \geq \overline{W}_L(\beta_{L,\ell}) \geq \overline{W}(\beta)$. Thus

$$\overline{X}(\beta) \geq E(\pi^o, \alpha^o, w^o) \geq \mu_0(1-\delta) + \delta\overline{W}(\beta) \geq \overline{X}(0).$$

If $\beta \in [\beta_{23}, 1]$it is optimal for the sender to send message H when $\theta = \ell$ and thus $v_\ell \geq \delta w_{L,\ell} = \delta\overline{W}(1)$. Therefore

$$E(\pi^o, \alpha^o, w^o) \geq \mu_0[(1-\delta) + \delta w^o_{H,h}] + (1-\mu_0)\delta\overline{W}(1).$$

Since $w^o_{H,h} = \overline{W}(\beta)$,

$$\overline{X}(\beta) \geq E(\pi^o, \alpha^o, w^o) \geq \mu_0[(1-\delta) + \delta\overline{W}(\beta)] + (1-\mu_0)\delta\overline{W}(1) \geq \overline{X}(0). \qquad \square$$

Lemma 8. Let (π, α, w) be an optimal tuple for W at some $\beta \in (0, 1]$. Then $\pi(H|\ell) \leq \bar{\pi}(\beta)$ and $\alpha_H(H) = 1$.

Proof: Let that (π, α, w) be an admissible for W at β with $\pi(H|\ell) > \bar{\pi}(\beta)$. Then, for any value of $\pi(H|h) \in [0, 1]$, the receiver's posterior that $\theta = h$ after receiving message H is strictly less than $1/2$. Hence, $\alpha_H(H) = 0$. If $\pi(H|h) = 1$, then $\beta_{H,h} = \beta$, $\beta_{H,\ell} < \beta$ (since $\bar{\pi}(\beta) \geq \pi_B$), and

$$E(\pi, \alpha, w) = \mu_0\delta w_{H,h} + (1-\mu_0)\delta w_{H,\ell} \leq \delta\overline{W}(\beta).$$

If $\pi(H|h) < 1$, then $\beta_{L,h} = 0$ and it is optimal for the sender to send message L after observing $\theta = h$, so

$$E(\pi, \alpha, w) = \mu_0 \delta w_{L,h} + (1 - \mu_0)\delta w_{H,\ell} \leq \delta \overline{W}(\beta).$$

But, $\beta_{L,\ell}(\beta, \pi^o(H|\ell)) > \beta$, and

$$E(\pi^o, \alpha^o, w^o) = \mu_0(1 - \delta) + \delta[\mu_0 \overline{W}(\beta) + (1 - \mu_0)\overline{W}(\beta_{L,\ell}] \geq \mu_0(1 - \delta) + \delta\overline{W}(\beta),$$

so $E(\pi^o, \alpha^o, w^o) > E(\pi, \alpha, w)$. Hence, no such admissible tuple (π, α, w) can be optimal.

Now, assume that (π, α, w) is optimal for W at β so that $\pi(H|\ell) \leq \bar{\pi}(\beta)$. Assume that $\alpha_H(H) < 1$. Then

$$E(\pi, \alpha, w) \leq \mu_0 \alpha_H(H)(1 - \delta) + \delta\overline{W}(\beta) < E(\pi^o, \alpha^o, w^o),$$

which is a contradiction. □

Lemma 9. Let (π, α, w) be an optimal tuple for W at some $\beta \in (0, 1)$. Then $\pi(H|h) = 1$ and $\alpha_L(L) = 1$.

Proof: By contradiction, assume that $\pi(H|h) < 1$. Then, $\pi(L|h) > 0$ and for the tuple to be adimissible, it must optimal for the sender to recommend L when $\theta = h$. But $\beta_{L,h} = 0$ and $w_{L,h} \leq \overline{W}(0)$ (since $\pi_B(L|h) = 0$). Therefore, the sender's expected continuation value after observing $\theta = h$ is bounded above by $\delta\overline{W}(0)$. Consider the modified tuple $(\hat{\pi}, \alpha, \hat{w})$ where $\hat{\pi}(m|\ell) = \pi(m|\ell)$ for $m = H, L$, $\hat{\pi}(H, h) = 1$, $\hat{\pi}(L|h) = 0$, $\hat{w}_{m,\ell} = w_{m,\ell}$ for $m = H, L$, $\hat{w}_{H,h} = \overline{W}(\beta) \geq \mu_0$, and $\hat{w}_{L,h} = 0$. Now $\beta_{H,h} = \beta$ and $\hat{w}_{H,h} \in W(\beta_{H,h})$. By previous lemma, $\alpha_H(H) = 1$, so it must be that the receiver's posterior that $\theta = h$ after receiving message H is greater or equal to $1/2$. After increasing $\pi(H|h)$ to $\hat{\pi}(H|h) = 1$ this posterior increases and hence $\alpha_H(H) = 1$ remains optimal for the receiver. One can readily verify that this new tuple is also admissible for W at β and the corresponding sender's expected continuation value after observing $\theta = h$ is $\alpha_H(H)(1 - \delta) + \delta\overline{W}(\beta) > \delta\mu_0$. Thus $E(\hat{\pi}, \alpha, \hat{w}) > E(\pi, \alpha, w)$, which is a contradiction.

Since $\pi(H|h) = \pi_B(H|h) = 1$, the receiver's posterior that $\omega = \ell$ after receiving the message $m = L$ is $\mathbb{P}[\ell|L] = 1$ and his unique best reply is to take action L, therefore $\alpha_L(L) = 1$. □

Lemma 10. For every $\beta \in (0, 1]$ there is an optimal tuple (π, α, w) such that $\pi(H|\ell) \geq \pi_B$.

Proof: Assume that (π, α, w) is an admissible tuple such that $\pi(H|\ell) < \pi_B$. Then $\pi(L|\ell) = 1 - \pi(H|\ell) \geq 1 - \pi_B = \pi_B(L|\ell) > 0$ and $\beta_{L,\ell} \leq \beta$. Since $\pi(L|\ell) > 0$, the tuple is admissible only if it is optimal for the sender to recommend L after $\theta = \ell$.

Hence $v_\ell \le \delta\overline{W}(\beta_{L,\ell}) \le \delta\overline{W}(\beta)$. Also $\beta_{H,h} = \beta$ (since $\pi(H|h) = \pi_B(H|h) = 1$), so $v_h \le \alpha_H(H)(1-\delta) + \delta\overline{W}(\beta) = (1-\delta) + \delta\overline{W}(\beta)$. Thus,

$$E(\pi, \alpha, w) \le \mu_0(1-\delta) + \delta\overline{W}(\beta).$$

Let (π^o, α^o, w^o) be the corresponding tuple for β. If $\beta \in (\beta_{23}, 1]$, then by previous lemma $E(\pi^o, \alpha^o, w^o) > \mu_0[(1-\delta) + \delta\overline{W}(\beta)] + (1-\mu_0)\delta\overline{W}(1) \ge \mu_0(1-\delta) + \delta\overline{W}(\beta)$. Therefore the tuple (π, α, w) cannot be optimal.

If $\beta \in (0, \beta_{23}]$, then by previous lemma, $\overline{X}(\beta) \ge E(\pi^o, \alpha^o, w^o) \ge \mu_0(1-\delta) + \delta\overline{W}(\beta)$. Hence, (π, α, w) is optimal if and only if

$$\overline{X}(\beta) = E(\pi, \alpha, w) = \mu_0(1-\delta) + \delta\overline{W}(\beta) = E(\pi^o, \alpha^o, w^o).$$

Thus, (π^o, α^o, w^o) is optimal and by definition $\pi^o(H|\ell) \ge \pi_B$. We note that this can happen only if $\overline{W}(\beta') = \overline{W}(\beta)$ for all $\beta' \in [\beta, \beta_{L,\ell}(\beta, \pi^o(H|\ell))$. □

Lemma 11. [Exploitation] For any $\beta \in [\beta_{23}, 1]$ the corresponding (π^o, α^o, w^o) is an optimal tuple for W at β. Moreover $\pi^o(H|\ell) = 1$.

Proof: If (π, α, w) is an admissible tuple at β with $\pi(H|\ell) < 1$, then $\beta_{L,\ell} < 1$ and $w_{L,\ell} \le \overline{W}(\beta_{L,\ell}) \le \overline{W}(1)$. Therefore

$$E(\pi, \alpha, w) \le \mu_0[(1-\delta) + \delta\overline{W}(\beta)] + (1-\mu_0)\delta\overline{W}(1).$$

But

$$E(\pi^o, \alpha^o, w^o) = \mu_0[(1-\delta) + \delta\overline{W}(\beta)] + (1-\mu_0)[(1-\delta) + \delta\overline{W}(\beta_{H,\ell}(\beta, 1)) > E(\pi, \alpha, w).$$

Thus (π, α, w) is not optimal. That is, an optimal tuple must have $\pi(H|\ell) = 1$.

Since $\bar{\pi}(\beta) \ge 1$, $\alpha^o_H(H) = 1$ is optimal for the receiver (for any $\pi^o(H|\ell)$), and given $\alpha^o_H(H) = 1$, the sender's expected payoff is maximized when $\pi^o(H|\ell) = 1$, $\pi^0(H|h) = 1$ and

$$w^o_{H,h} = \overline{W}(\beta), \quad w^o_{H,\ell} = \overline{W}(\beta_{H,\ell}(\beta, 1)).$$

Given $\pi^o(H|\ell) = 1$ and that $\pi_B(H|h) = 1$, $\beta_{L,h}$ is not well defined by Bayes' rule; by continuity we let $\beta_{L,h} = 0$ (as this is the posterior for any $\pi^o(H|h) < 1$) and for convenience choose $w_{H,h} = 0 \in W(0)$. □

Lemma 12. Let (π, α, w) be an admissible tuple for some $\beta \in [0, \beta_{23})$ such that $\pi(H|\ell) = 1$. Then (π, α, w) is not optimal.

Proof: By contradiction, assume that (π, α, w) is optimal. Then, $\pi(H|\ell) \le \bar{\pi}(\beta)$ and $\pi(H|\ell) = \hat{\pi}(\beta) = 1$. Therefore $\beta \le \beta_{23}$, a contradiction. □

Lemma 13. [Indifference] For any $\beta \in [0, \beta_{23})$ the corresponding (π^o, α^o, w^o) is an optimal tuple for W at β. Moreover $\pi_B \leq \pi^o(H|\ell) < 1$ and the sender is indifferent between messages H and L when $\theta = \ell$.

Proof: Since $X = B(W)$ has a compact graph, for each $\beta \in [0, 1]$ there is an optimal policy. Fix $\beta \in [0, \beta_{23})$ and let (π, α, w) be an optimal policy for W at β and (π^o, α^o, w^o) be the corresponding admissible tuple. By Lemma 8, $\alpha_H(H) = 1$ and $\pi(H|\ell) \leq \bar{\pi}(\beta)$. By Lemmas 10 and 10, $\pi(H|\ell) < 1$ and without loss of generality, that $\pi(H|\ell) \geq \pi_B$. Assume by contradiction that $E(\pi, \alpha, w) > E(\pi^o, \alpha^o, w^o)$. Since $\pi(H|\ell) < 1$, sending message L after observing $\theta = \ell$ must be optimal for the sender and thus $(1 - \delta) + \delta w_{H,\ell} \leq \delta w_{L,\ell}$ and

$$E(\pi, \alpha, w) = \mu_0(1 - \delta) + \delta[\mu_0 w_{H,h} + (1 - \mu_0) w_{L,\ell}],$$

where $w_{H,h} \leq \overline{W}(\beta)$, $w_{H,\ell} \leq \overline{W}(\beta_{H,\ell})$ and $w_{L,\ell} \leq \overline{W}(\beta_{L,\ell})$. Increasing $w_{H,h}$ strengthens the incentives for sending message H when $\theta = h$, and thus optimally $w_{H,h} = \overline{W}(\beta)$. We also have that

$$E(\pi^o, \alpha^o, w^o) = \mu_0(1 - \delta) + \delta[\mu_0 \overline{W}(\beta) + (1 - \mu_0) w^o_{L,\ell}],$$

where $w^o_{L,\ell} \in [\overline{W}_L(\beta^o_{H,\ell}), \overline{W}(\beta^o_{H,\ell})]$. Therefore $w_{L,\ell} > w^o_{L,\ell}$, which implies that $\beta_{L,\ell} > \beta^o_{L,\ell}$. Hence, $\pi(H|\ell) > \pi^o(H|\ell)$ and $\beta_{H,\ell} < \beta^o_{H,\ell}$, so $w_{H,\ell} \leq w^o_{H,\ell}$. By definition, $(1 - \delta) + \delta w^o_{H,\ell} = \delta w^o_{L,\ell}$, which implies that $(1 - \delta) + \delta w_{H,\ell} < \delta w_{L,\ell}$. But then, admissibility requires that $\pi(H|\ell) = 0$, a contradiction. □

Define $\mathcal{C}_0^{++} = \{W \in \mathcal{C}_0 \mid \overline{W}(\beta) \text{ is strictly increasing in } \beta\}$.

Lemma 14. Assume $W \in \mathcal{C}_0^{++}$. Then $\overline{X}(\beta)$ is strictly increasing in $\beta \in [0, 1]$.

Proof: Let $0 < \beta < \hat{\beta}$. We now show that $\overline{X}(\beta) < \overline{X}(\hat{\beta})$. Let (π^o, α^o, w^o) and $(\hat{\pi}^o, \hat{\alpha}^o, \hat{w}^o)$ be the corresponding optimal tuples. Then $\hat{\beta}_{L,\ell} \equiv \beta_{L,\ell}(\hat{\beta}, \hat{\pi}^o) \geq \beta_{L,\ell}(\beta, \pi^o) \equiv \beta_{L,\ell}$. By contradiction, assume that $\hat{\beta}_{L,\ell} < \beta_{L,\ell}$. Then $\hat{\pi}^o(H|\ell) < \pi^o(H|\ell) \leq \bar{\pi}(\beta) < \bar{\pi}(\hat{\beta})$, so $\hat{\beta}_{H,\ell} > \beta_{H,\ell}$ and

$$\overline{W}(\hat{\beta}_{L,\ell}) - \overline{W}_L(\hat{\beta}_{H,\ell}) < \overline{W}(\beta_{L,\ell}) - \overline{W}(\beta_{H,\ell}) \leq (1 - \delta)/\delta,$$

contradicting the definition of $\hat{\pi}(H|\ell)$ (by the previous inequality, $\hat{\pi}(H|\ell)$ should be increased to increase $\hat{w}_{L,h}$ and improve $E(\hat{\pi}^o, \hat{\alpha}^o, \hat{w}^o)$). Note that even if $\overline{W}(\beta_{L,\ell}) - \overline{W}_L(\beta_{H,\ell}) > (1 - \delta)/\delta$, $\overline{W}(\beta_{L,\ell}) - \overline{W}(\beta_{H,\ell}) \leq (1 - \delta)/\delta$ by the definition of $\bar{\pi}(\beta)$.

Previously we also established that $\overline{X}(\beta) \geq \overline{X}(0)$ for all $\beta > 0$. Thus, $\overline{X}(\beta)$ is strictly increasing in $\beta \in [0, 1]$. □

Proposition 2. [Monotonicity] $\overline{V}(\beta)$ is weakly increasing in β.

Proof: Construct the equilibrium correspondence V starting with the initial seed

$$W^0 = \{(\beta, w) \mid \beta \in [0,1] \text{ and } w \in [0, 1+\gamma\beta] \text{ for each } \beta \in [0,1]\},$$

where $\gamma > 0$ (small). Clearly $W^0 \in \mathcal{C}_0^{++}$. Let $W^{k+1} = \tilde{B}(W^k)$ for $k = 0, 1, \dots$ Then, $W^k \in \mathcal{C}_0^{++}$ for each $k \geq 1$. Since $V = \lim_{k\to\infty} W^k$, we have that $\overline{V}(\beta)$ is weakly increasing in β. □

B. Equilibrium Behavior and Behavioral Convergence: Proofs

Lemma 15. For $\beta \in [0, \beta_{23})$, $\beta_{H,\ell}(\beta, \pi(\beta))$ and $\beta_{L,\ell}(\beta, \pi(\beta))$ are increasing functions of β.

Proof: Let $\beta \in [0, \beta_{23})$ 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 $\beta \in [0, \beta_{23})$ 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 $\beta_{L,\ell}(\beta, \pi(\beta)) < \beta_{L,\ell}(\beta', \pi(\beta'))$. By definition of $\pi(\beta)$,

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

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

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

by definition of $\pi(\beta')$ it must be that $\overline{V}(\beta_{L,\ell}(\beta', \pi(\beta'))) > \overline{V}(\beta_{L,\ell}(\beta, \pi(\beta)))$ and $\beta_{L,\ell}(\beta', \pi(\beta')) > \beta_{L,\ell}(\beta, \pi(\beta))$. □

Recall that for each reputation β, $\pi(\beta)$ is the sender's strategy 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 6.2 considers the reputation maintenance strategy based at β_0. By Lemma 15, $\beta_H \leq \beta_{H,\ell}(\beta, \pi(\beta)) < \beta$ 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), 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 16. $V^\pi(\beta, \delta) \geq \overline{V}(\beta, \delta)$ for all $\beta \in [\beta_H, \beta_L]$. Conversely, if $[\beta_0, \beta_L]$ is in Region 2 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

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

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

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

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 $[\beta_0, \beta_L]$ is in Region 2. Then $w_{H,\ell} = \overline{V}(\beta_{H,\ell}, \delta)$ and $\overline{V}(\beta, \delta) = T^\pi(\overline{V}(\cdot, \delta))(\beta)$ for all $\beta \in [\beta_H, \beta_L]$. That is $\overline{V}(\cdot, \delta)$ is a fixed point of T^π, and therefore $\overline{V} = V^\pi$. □

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

Lemma 17 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 17 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 abstract from the complexities involved in the proof of Lemma 18.

Lemma 17. (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 2. Then there exists a constant $D > 0$ (independent of δ and π_0) such that

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

Proof: This is a Corollary of Lemma 18. 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

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

By Bayes' 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 $\hat{\beta} \in [\beta_H, \beta_L]$,

$$\frac{\pi_0}{\pi_B} = L_{1,0} \ge L_{n,m} \ge L_{0,1} = \frac{1-\pi_0}{1-\pi_B} \quad \Longleftrightarrow \quad \frac{n-1}{m} \le R \le \frac{n}{m-1}$$

$$\Longleftrightarrow \quad \frac{n-1}{n+m-1} \le \frac{R}{R+1} \le \frac{n}{n+m-1}. \qquad (**)$$

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. 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 18. 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:

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

Note that

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

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

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

By Taylor series expansion

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

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 18,

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

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 1: 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

$$\overline{V}(\beta_{23}, \delta) - (1-\epsilon) = \frac{\delta - \mu_0(2\delta - 1)}{1 - \delta\mu_0} - (1-\epsilon) > \epsilon/2.$$

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 $v_0 = \overline{V}(\epsilon, \delta) \leq 1 - \epsilon < \frac{\delta - \mu_0(2\delta-1)}{1-\delta\mu_0} = \overline{V}(\beta_{23}, \delta)$, we have that $\beta_0 < \beta_{23}$. For as long as $\beta_{k-1} < \beta_{23}$, 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 $v_k < \overline{V}(\beta_{23}, \delta)$, and hence $\beta_k < \beta_{23}$, 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 20 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 19. 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 $k = 1, \ldots, \tilde{K}$ 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. □

Lemma 20. Let $\hat{\beta}$ be such that $\overline{V}(\hat{\beta}) = V^{1,2}(\pi_B, \delta)$. Then

1. there is $\underline{\beta} \in [\hat{\beta}/\lambda_1, \lambda_1\hat{\beta}]$ such that $\overline{V}(\lambda_1\underline{\beta}) - \overline{V}(\lambda_0\underline{\beta}) = (1-\delta)/\delta$;
2. 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 $\beta < \underline{\beta}$;
3. 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 19). 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. □

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 1, 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 18,

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

for all $\beta \in [0, \hat{\epsilon}]$. By Lemma 16, 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

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

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

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

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

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

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 $K\Delta_j + E^{n\times 1}\eta \le 2\hat{\epsilon}$ for all $\delta_j \in [\bar{\delta},1)$. Let $\delta_j \in [\bar{\delta},1)$. Then

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

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

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

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

If $\pi_{2K-1} > \pi^{n\times 1}$, stop and make $M = 2$. Otherwise, Corollary 2 again implies that $\pi_{2K} \ge \pi_K + \eta$. 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 $\pi(\beta^{n\times 1}) \ge \pi(\beta_{MK-1}) > \pi^{n\times 1}$ 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, $k=2,\ldots,n+1$. 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 $\tilde\pi_k = \pi(\tilde\beta_k)$, $k=0,\ldots,n+1$. 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 16, 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

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

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 1, 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 2: 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 $k \le \bar{k}$ 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 20 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 20. 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.

Aldrich, Eric M., Jesús Fernández-Villaverde, A. Ronald Gallant and Juan F. Rubio-Ramirez (2011), "Tapping the supercomputer under your desk: Solving dynamic equilibrium models with graphics processors," *Journal of Economic Dynamics and Control*, **35**(3), 386–393.

Andreasen, Martin (2012), "On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-Linear DSGE Models," *Review of Economic Dynamics*, **15**(3), 295–316.

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 (2023), "Persuasion for the Long Run," Forthcoming at the *Journal of Political Economy.*

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.

Fernández-Villaverde, Jesús and Oren Levintal (2018), "Solution methods for models with rare disasters," *Quantitave Economics*, **9**(2), 903–944.

Fernández-Villaverde, Jesús and David Zarruk Valencia (2018), "A Practical guide to parallelization in economics", Working Paper.

Fudenberg, Drew, Ying Gao and Harry Pei (2022), "A reputation for honesty", *Journal of Economic Theory*, **204**.

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.

Gourio, Francois (2012), "Disaster risk and business cycles", *American Economic Review*, **102**(6), 2734-2766.

Gourio, Francois (2013), "Credit Risk and Disaster Risk", *American Economic Journal: Macroeconomics*, **5**(3), 1-34.

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.

Isoré Marlène and Urszula Szczerbowicz (2017), "Disaster risk and preference shifts in a New Keynesian model," *Journal of Economic Dynamics and Control*, **79**, 97–125.

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 (2023), "Repeated communication with private lying costs," *Journal of Economic Theory*, Elsevier, vol. 210(C).

Petrosky-Nadeau, Nicolas, Lu Zhang and Lars-Alexander Kuehn (2018), "Endogenous Disasters," *American Economic Review*, **108**(8), 2212–2245.

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.