DYNAMIC CHEAP TALK WITHOUT FEEDBACK

Abstract. We study a dynamic sender-receiver game in which the state evolves according to a Markov chain observed by the sender, who does not observe the receiver's action. Despite the absence of feedback, dynamic interaction remains valuable. We show that long-run interaction partially restores commitment: any equilibrium payoff of a persuasion model with partial commitment, where the sender can deviate only to signaling policies that preserve the marginal distribution over messages, as well as any convex combination of such payoffs across distributions over messages, can be achieved as a uniform equilibrium payoff in the dynamic game. When the sender's payoff is state-independent, she fully bridges the commitment gap and achieves the Bayesian persuasion payoff.

1. Introduction

Dynamic interactions rely on feedback. An expert with valuable information has to decide when and how to reveal information to a decision maker. The expert provides information conditional on the decision maker's good behaviour but can punish him by not providing further information when he deviates (carrot-and-stick policy). However, in many environments, such feedback is absent: the expert does not observe the decision maker's actions. Can dynamic interaction still be valuable in such settings? If so, what are the limits of what can be achieved without feedback?

We consider an infinite-horizon dynamic cheap talk model. In each period, the sender privately observes the current state and sends a message to the receiver, who then chooses an action. The action is not disclosed to the sender (i.e., there is no feedback). The players' payoffs depend only on the current state and the receiver's action, and the state evolves over time according to a Markov chain.

Despite the absence of feedback, dynamic interaction remains valuable. The key observation is that, although the sender does not observe the receiver's actions, the receiver observes the entire history of messages. This allows him to monitor the sender's behavior over time using using statistical tests based on message frequencies. A single message cannot reveal whether the sender has deviated. However, over time, the receiver can collect data on the messages and check whether their distribution matches what is prescribed. If the sender deviates in a detectable way, the receiver can respond by no longer following her recommendations. As a result, if the prescribed strategy is optimal for the

---

Date: March 28, 2026.

Department of Economics, University of Bonn. Email: ajain@uni-bonn.de.

I am indebted to Tristan Tomala and Nicolas Vieille for their advice and encouragement throughout the project.

sender among all undetectable deviations—those that preserve the empirical distribution over messages—outcomes beyond those achievable in the one-shot game can be achieved.

Our main result (Theorem 1) shows that any equilibrium payoff of the persuasion model with partial commitment—where the sender can deviate only to signaling policies that preserve a fixed distribution over messages—can be achieved as a uniform equilibrium payoff in the dynamic game. Moreover, any convex combination of such payoffs across message distributions can also be achieved. This static persuasion model is closely related to the model of Lin and Liu (2024). We show that it is without loss to focus on direct signaling policies in which the sender recommends an action to the receiver (Proposition 1). In this formulation, equilibrium is characterized by the receiver following the recommended action and by the sender having no profitable deviations among policies that preserve the marginal distribution over messages, where the prior is the invariant distribution of the Markov chain.

Thus, dynamic communication without feedback partially restores commitment, improving the sender's payoff relative to the one-shot cheap talk game. In particular, when the sender's payoff is state-independent, the sender can fully bridge the commitment gap and achieve the Bayesian persuasion payoff, which is the maximal payoff attainable even under full commitment.

The intuition is as follows. The sender induces the same joint distribution over states and recommendations in every period (in expectation), which is feasible under a pseudo-renewal Markov chain. As a result, the empirical distribution over messages converges to a fixed target that the receiver can monitor, making any deviation that alters it detectable. Conditional on each recommendation, the induced beliefs match the desired posterior, so the receiver optimally follows. Thus, the sender is effectively constrained to deviations that preserve message frequencies, as in the static persuasion model with partial commitment, and convex combinations are obtained by varying these frequencies across blocks.

The previous result identifies a natural subset of uniform equilibrium payoffs. We next examine how tight this characterization is. Example 2 shows that, in general, uniform equilibrium payoffs need not lie in the convex hull of partial commitment payoffs. Proposition 2 provides a necessary condition: equilibrium outcomes must be robust to a class of undetectable deviations based on permutations of the state process. These deviations preserve the statistical properties of the process and are therefore indistinguishable from equilibrium play, following the approach of Renault, Solan, and Vieille (2013). Finally, Proposition 3 shows that when states are i.i.d., this restriction becomes tight: all deviations that preserve message frequencies are feasible and undetectable, and the static benchmark—the convex hull of equilibrium payoffs of the persuasion problem with partial commitment—becomes exact. An implication is that this provides a micro-foundation for the persuasion model with partial commitment.

Related Literature: Our paper relates to the growing literature on strategic communication in dynamic environments where uncertainty evolves. For example, Renault, Solan, and Vieille (2013), Escobar and Toikka (2013), Renault, Solan, and Vieille (2017), Margaria and Smolin (2018) and Best and Quigley (2024).

In particular, our paper is closely related to Kuvalekar, Lipnowski, and Ramos (2022), who also study a repeated cheap talk model without feedback. They consider an i.i.d. environment and characterize Perfect Bayesian equilibrium payoffs for a fixed discount factor via an equivalence to static cheap talk with money burning. In contrast, we consider a Markovian setting and study uniform equilibrium payoffs. This leads to different conclusions: for example, they show that the sender cannot achieve the Bayesian persuasion payoff even with very patient players, whereas in our setting she can. In particular, when the sender's preferences are state-independent, they show that repetition does not improve the sender's payoff at all, while in precisely this case we show that she can achieve the Bayesian persuasion payoff.

Our equilibrium construction is based on quota mechanisms, where agents are required to match prescribed message frequencies; see, for example, Jackson and Sonnenschein (2007), Escobar and Toikka (2013) and Frankel (2016).

In particular, our approach is closest to Renault, Solan, and Vieille (2013), where the receiver monitors message frequencies using quota-based tests. In their setting, the set of undetectable deviations coincides exactly with state permutations, and the sender can punish deviations by withholding information. In contrast, in our setting the sender does not observe the receivers actions, which limits her ability to discipline him. This leads the sender to provide action recommendations rather than report states; as a result, state permutations remain a natural class of undetectable deviations, but are no longer sufficient.

Our main result links the dynamic game to a static sender-receiver problem, connecting our analysis to the literature on communication with partial commitment. In cheap talk (Crawford and Sobel, 1982; Green and Stokey, 2007), the sender cannot commit, while in Bayesian persuasion (Kamenica and Gentzkow, 2011), she fully commits to a signaling policy. Our static benchmark lies in between: the sender can deviate to any policy that preserves the marginal distribution over messages. This restriction is very closely related to the notion of credibility in Lin and Liu (2024). The key difference is that in their model the sender chooses the marginal distribution over messages, whereas in our setting it is fixed ex ante. Moreover, in our dynamic environment players can achieve convex combinations of such outcomes through long-run interaction.

More broadly, our static formulation relates to work that weakens commitment in Bayesian persuasion, for example through noisy deviations or information constraints;

see, for instance, Le Treust and Tomala (2019); Min (2021); Lipnowski, Ravid, and Shishkin (2022). Our contribution is to show that long-run interaction can substitute for commitment, albeit only partially.

The rest of the paper is organized as follows. In section 2, we introduce the dynamic cheap talk model with no feedback and discuss the value of commitment. In section 3, we present the persuasion model with partial commitment and provide our main results. In section 4, we conclude and discuss future work. All omitted proofs are in appendix A.

2. Model

We consider a dynamic game between two players: the sender (she) and the receiver (he). The state of the world evolves over time in a discrete-timed infinite horizon. At each period $n \in \mathbb{N}$, the sender observes the current state $\omega_n \in \Omega$ and sends a message $m_n \in M$ to the receiver. Upon observing the message m_n, the receiver takes an action $a_n \in A$. The action a_n is not disclosed to the sender. The sender's and the receiver's payoff in the period equals $u_S(\omega_n, a_n)$ and $u_R(\omega_n, a_n)$ respectively.[1]

The state evolves according to a Markov chain. Given the current state ω_n, the state in the next period ω_{n+1} is drawn according to the transition matrix $Q(\omega_{n+1} \mid \omega_n)$. We assume the Markov chain is irreducible, aperiodic and pseudo-renewal.

Definition 1. A Markov chain $\{\omega_n\}_{n\in\mathbb{N}}$ is **pseudo-renewal** if $\omega \neq \tilde{\omega}$, then $Q(\omega \mid \tilde{\omega}) = \alpha_\omega$ for some $(\alpha_\omega)_\omega \geq 0$.

A Markov chain is pseudo-renewal if, for any two distinct states, the transition probability depends only on the destination state and not on the current state. For example, this condition holds for the case of i.i.d chains or for any chain with binary states.[2]

Let $\mu \in \Delta(\Omega)$ denote the unique invariant distribution of the Markov chain.[3] The initial state ω_1 is drawn according to μ. The sets of states Ω, actions A, and messages M are all finite. We assume that $|M| \geq |A|$.

The players' strategies map their information into choices. For the sender, a strategy is a mapping $\sigma_n : (\Omega \times M)^{n-1} \times \Omega \to \Delta(M)$, which specifies a distribution over messages given past states and messages and the current state. Without loss of generality, we restrict attention to strategies that depend only on the current state and past messages.[4] For the receiver, a strategy is a mapping $\tau_n : (M \times A)^{n-1} \times M \to \Delta(A)$, which specifies a distribution over actions given past messages and actions and the current message. Given

[1]Neither player observes realized payoffs. In particular, the sender does not observe the action a_n, while the receiver does not observe the state ω_n.

[2]Renault, Solan, and Vieille (2013, 2017) and Hörner, Mu, and Vieille (2017) also analyze dynamic sender-receiver games, where the Markov chain is pseudo-renewal.

[3]The invariant distribution satisfies $\sum_{\omega\in\Omega} \mu(\omega)Q(\omega' \mid \omega) = \mu(\omega')$ for all $\omega' \in \Omega$.

[4]The sender does not gain from conditioning on past states because (i) the receiver does not observe states and his strategy depends only on messages, and (ii) future states depend only on the current state.

a strategy profile (σ, τ), the δ-discounted payoff of player $i \in \{S, R\}$ is

$$\gamma_i^\delta(\sigma, \tau) = \mathbb{E}_{\sigma,\tau}\Big[(1-\delta)\sum_{n=1}^{\infty} \delta^{n-1} u_i(\omega_n, a_n)\Big].$$

Uniform equilibrium: In the dynamic game, players do not observe their realized payoffs, and equilibrium strategies typically vary with the discount factor. We focus on sufficiently patient players and allow for approximate equilibria. We therefore adopt uniform equilibrium as the solution concept: a payoff vector can be approximated by equilibria of the discounted game for all sufficiently large discount factors.

Definition 2. A payoff vector $\boldsymbol{\gamma} = (\gamma_S, \gamma_R) \in \mathbb{R}^2$ is a **uniform equilibrium payoff** if for every $\epsilon > 0$, there exists a $\delta_0 \in (0, 1)$ and a strategy profile (σ^*, τ^*) such that for all $\delta \geq \delta_0$,

$$\gamma_S^\delta(\sigma^*, \tau^*) + \epsilon \geq \gamma_S \geq \gamma_S^\delta(\tilde{\sigma}, \tau^*) - \epsilon \quad \forall \tilde{\sigma}, \tag{2.1}$$

$$\gamma_R^\delta(\sigma^*, \tau^*) + \epsilon \geq \gamma_R \geq \gamma_R^\delta(\sigma^*, \tilde{\tau}) - \epsilon \quad \forall \tilde{\tau}. \tag{2.2}$$

A payoff vector $\boldsymbol{\gamma}$ is a uniform equilibrium payoff if for every $\epsilon > 0$, there exists $\delta_0 \in (0, 1)$ and a strategy profile (σ^*, τ^*) such that for all $\delta \geq \delta_0$, it is an ϵ-equilibrium and yields a payoff within ϵ of $\boldsymbol{\gamma}$. Denote by $\mathcal{U}$ the set of all uniform equilibrium payoffs.

Value of Commitment: We show that the Bayesian persuasion payoff evaluated at the invariant distribution is an upper bound on the sender's equilibrium payoff. To establish this result, we adopt a belief-based representation of the dynamic game, under which payoffs can be tracked through the receiver's beliefs.

In the dynamic game, the receiver updates his belief over the states based on sender's message. Since the state follows a Markov chain, beliefs evolve across periods. Let q_t and p_t denote the receiver's belief about the state ω_t before and after observing the message at stage t, respectively. The transition matrix Q links the two beliefs:

$$q_{t+1} = p_t Q.$$

Given the sender's strategy σ, in any ϵ-equilibrium the receiver plays a best response to his posterior belief in almost all periods. Since actions are unobserved, any deviation to a myopic best response is profitable. Therefore, equilibrium payoffs can be tracked using the receiver's beliefs alone.

For $i \in \{S, R\}$, define the indirect utility

$$\hat{u}_i(p) := \mathbb{E}_p[u_i(\omega, a^*(p))],$$

where $a^*(p)$ is the receiver's best response given belief p.[5]

[5]If multiple actions are optimal for the receiver, we select the one preferred by the sender. Formally, let $A^*(p) := \operatorname{argmax}_{a\in A} \mathbb{E}_p[u_R(\omega, a)]$, and define $a^*(p) := \operatorname{argmax}_{a\in A^*(p)} \mathbb{E}_p[u_S(\omega, a)]$.

In any period, the sender's strategy induces a Bayes-plausible distribution over posterior beliefs, i.e., the expected posterior equals the prior. In the static persuasion model, Kamenica and Gentzkow (2011) show that the sender's maximal payoff, given prior π, is the concave envelope of the sender's indirect utility, denoted by $\operatorname{Cav} \hat{u}_S(p)$.

The same reasoning applies period by period in the dynamic game. In period n, the sender induces a Bayes-plausible distribution over posterior beliefs with prior q_n. Hence, her expected payoff in period n is at most $\operatorname{Cav} \hat{u}_S(q_n)$. Therefore,

$$\gamma_S^\delta \leq (1-\delta) \sum_{n=1}^{\infty} \delta^{n-1} \mathbb{E}_\sigma[\operatorname{Cav} \hat{u}_S(q_n)] \tag{2.3}$$

$$\leq (1-\delta) \sum_{n=1}^{\infty} \delta^{n-1} \operatorname{Cav} \hat{u}_S(\mathbb{E}_\sigma[q_n]) \tag{2.4}$$

$$= \operatorname{Cav} \hat{u}_S(\mu), \tag{2.5}$$

where the second inequality follows from Jensen's inequality, and the last equality uses $\mathbb{E}_\sigma[q_n] = \mu$ for all n. Thus, the Bayesian persuasion payoff at μ is an upper bound in the dynamic game.[6]

The stage game corresponds to cheap talk, where the sender cannot commit to how messages are generated. Without commitment, the sender's payoff is below the persuasion benchmark. When preferences are state independent, Lipnowski and Ravid (2020) show that the maximal payoff is given by the quasi-concave envelope of $\hat{u}_S$. In this case, the gap between the concave and quasi-concave envelopes captures the value of commitment (see Figure 1a).

Our objective is to characterize how much of this gap can be recovered through dynamic interaction without feedback. We show that the sender's equilibrium payoff improves upon cheap talk, and that when preferences are state independent, the sender attains the Bayesian persuasion payoff.

3. Main Results

Our main result partially characterizes the set of uniform equilibrium payoffs in terms of a static persuasion model with partial commitment. We first define the static model. We then show that any equilibrium payoff of the static model, and its convexification, can be sustained as a uniform equilibrium payoff. Finally, we show that there may exist equilibrium payoffs outside this set and provide a sufficient condition for equilibrium.

3.1. **Persuasion with partial commitment.** The key benchmark in our analysis is a static persuasion model with partial commitment, in which the sender chooses a signaling policy subject to a fixed marginal distribution over messages and can deviate only to policies that preserve this marginal. This formulation is closely related to the notion of

[6] Renault, Solan, and Vieille (2017) obtain the same upper bound even when the sender can commit against short-run receivers.

credibility in Lin and Liu (2024). The environment coincides with the stage game of the dynamic model. In particular, the sets of states Ω, messages M, and actions A, as well as the payoff functions u_S and u_R, are the same as in the dynamic game.

Fix a prior $\pi \in \Delta(\Omega)$ and a distribution over messages $q \in \Delta(M)$. The set of feasible signaling policies is

$$(3.1) \qquad \Sigma(\pi, \lambda) := \left\{ \rho : \Omega \to \Delta(M) \ : \ \sum_{\omega \in \Omega} \pi(\omega)\rho(m \mid \omega) = \lambda(m) \text{ for all } m \in M \right\}.$$

Consider a professor who assigns grades to a class of students. The set of grades is $M = \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}, \mathrm{E}\}$. The university imposes a fixed distribution over grades, for example $\lambda = \left(\frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{2}{10}, \frac{1}{10}\right)$. The professor observes each student's ability $\omega \in \Omega$ and assigns a grade according to a policy $\rho : \Omega \to \Delta(M)$. The constraint requires that the fraction of students receiving each grade matches λ. For instance, exactly 10% of students must receive an A, regardless of the distribution over abilities.

The game proceeds as follows. The sender chooses a signaling policy $\rho \in \Sigma(\pi, \lambda)$. The receiver observes the realized message $m \in M$ and chooses an action according to a response rule $\kappa : M \to \Delta(A)$. Given a strategy profile (ρ, κ), player i's expected payoff is

$$\sum_{\omega \in \Omega} \pi(\omega) \sum_{m \in M} \rho(m \mid \omega) \sum_{a \in A} \kappa(a \mid m)\, u_i(\omega, a), \quad \text{for } i \in \{S, R\}.$$

Let $\mathcal{P}(\pi, \lambda)$ denote the persuasion model with partial commitment with prior $\pi \in \Delta(\Omega)$ and distribution over messages $\lambda \in \Delta(M)$

An equilibrium is a strategy profile (ρ, κ) in which the receiver best responds to posterior beliefs induced by each message and the sender cannot profitably deviate within $\Sigma(\pi, \lambda)$. Let $\mathcal{E}(p, q)$ denote the set of equilibrium payoffs.

We say that a policy is a direct recommendation policy if messages correspond to action recommendations, i.e., $M \equiv A$. The following proposition shows that it is without loss to focus on such policies.

Proposition 1. *If $A \subseteq M$, then for any equilibrium (ρ, κ) of $\mathcal{P}(\pi, \lambda)$, there exists an equivalent equilibrium $(\tilde{\rho}, \tilde{\kappa})$ of $\mathcal{P}(\pi, \tilde{\lambda})$, where messages are action recommendations, the receiver obeys them, and the induced marginal distribution over actions is given by $\tilde{\lambda}(a) = \sum_{m \in M} \lambda(m)\kappa(a \mid m)$.*

Intuitively, each message matters only through the action it induces, so any equilibrium can be replaced by an equivalent one in which messages are action recommendations.[7] This observation is also made in Lin and Liu (2024, Proposition 1).

In view of the above result, we henceforth restrict attention to direct recommendation policies. Thus, $M \equiv A$, and the sender chooses a recommendation policy $\rho : \Omega \to \Delta(A)$, where $\lambda(a) = \sum_{\omega} \pi(\omega)\rho(a \mid \omega)$ for all $a \in A$. Any such policy induces an outcome

[7] Two strategy profiles are equivalent if they result in the same joint distribution over states and actions.

$\nu \in \Delta(\Omega \times A)$ given by

$$\nu(\omega, a) = \pi(\omega)\rho(a \mid \omega).$$

We will refer to ν as the induced outcome.

An outcome ν satisfies Bayes plausibility if its marginal over states equals π. The receiver's equilibrium condition reduces to obedience: for every $a \in A$,

$$\text{(Eq-R)} \qquad \sum_{\omega} \nu(\omega, a)\big(u_R(\omega, a) - u_R(\omega, b)\big) \geq 0 \quad \forall b \in A.$$

The sender's equilibrium condition is that no alternative outcome with the same marginals yields a higher payoff:[8]

$$\text{(Eq-S)} \qquad \sum_{\omega, a} \big[\nu(\omega, a) - \tilde{\nu}(\omega, a)\big] u_S(\omega, a) \geq 0$$

for all $\tilde{\nu} \in \Delta(\Omega \times A)$ such that $\sum_a \tilde{\nu}(\omega, a) = \pi(\omega)$ and $\sum_\omega \tilde{\nu}(\omega, a) = \lambda(a)$.

It is convenient to describe outcomes in terms of the posterior beliefs they induce. Let $\{p_a\}_{a \in A}$ denote the posterior belief following recommendation a, defined by

$$p_a(\omega) = \frac{\nu(\omega, a)}{\lambda(a)} \quad \text{whenever } \lambda(a) > 0.$$

These beliefs satisfy Bayes plausibility, $\sum_{a \in A} \lambda(a) p_a = \pi$.

Obedience requires that each recommended action a is optimal given p_a, i.e.,

$$a \in \arg\max_{b \in A} \sum_{\omega \in \Omega} p_a(\omega)\, u_R(\omega, b) \quad \forall a \in A.$$

The sender's payoff depends only on the induced distribution over posterior beliefs, and her equilibrium condition is that no deviation to another collection of posterior beliefs satisfying Bayes plausibility with weights λ yields a higher payoff.

This persuasion model with partial commitment lies between Bayesian persuasion and cheap talk. In Bayesian persuasion, the sender faces no restriction on the marginal distribution over messages and cannot deviate after choosing a policy, whereas in cheap talk the sender can deviate freely. Here, deviations are restricted to policies that preserve the marginal distribution over messages. This formulation is closely related to the credible persuasion framework of Lin and Liu (2024), except that we fix the marginal distribution λ ex ante rather than allowing the sender to choose it.

Given prior π, let $e^*(\pi, \lambda)$ denote the sender-optimal equilibrium payoff. Define

$$e^*(\pi) := \max_{\lambda \in \Delta(A)} e^*(\pi, \lambda),$$

the sender's maximal equilibrium payoff across all marginals.

Remark 1. If the sender's payoff is state independent, then

$$e^*(\pi) = \operatorname{Cav} \hat{u}_S(\pi).$$

[8]The receiver's obedience condition need not hold under the deviation $\tilde{\nu}$.

Deviations that preserve the marginal distribution do not change the induced actions, so the sender can achieve the Bayesian persuasion payoff.

To illustrate the persuasion model with partial commitment, consider the following example, adapted from Lipnowski and Ravid (2020).

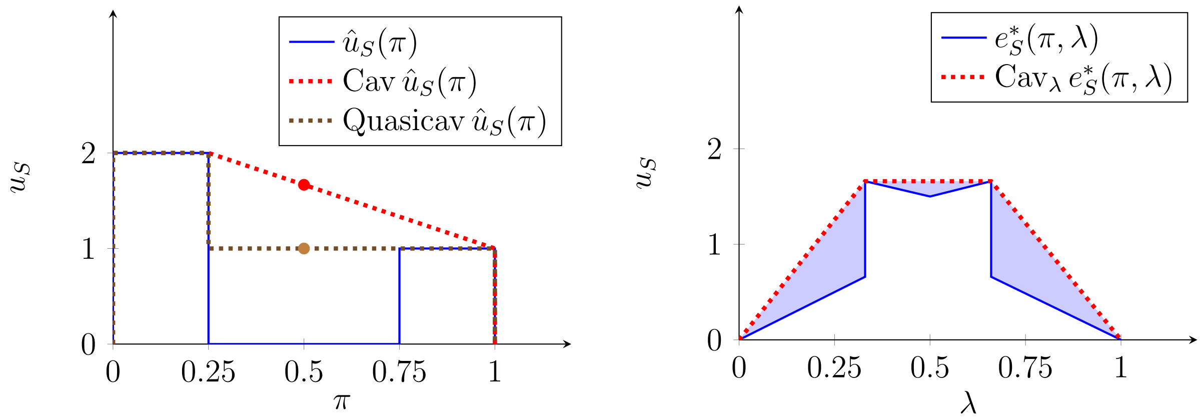
Example 1. *A political think tank advises a lawmaker on passing reforms. The lawmaker is contemplating whether to pass one of two possible reforms, denoted by 1 and 2, or to not pass any. These actions are denoted by* a_1 , a_2 *and* a_0 *respectively. The think tank preferences are state independent: it prefers reform 2 over reform 1 and having a reform over no reform. The lawmaker only wants to pass a law if he is sufficiently confident that the reform will be good. Consider the state space* $\Omega = \{\omega_1, \omega_2\}$*, where* ω_i *refers to the state where reform* i *is good. And let* $M \equiv A = \{a_0, a_1, a_2\}$ *denote the message space. The think tank and the law maker's preferences are captured by the following payoff matrix:*

	a_2	a_1	a_0
ω_1	$(2,0)$	$(1,4)$	$(0,3)$
ω_2	$(2,4)$	$(1,0)$	$(0,3)$

The receiver's optimal action, as a function of $p = \mathbb{P}(\omega_1)$*, is*

$$a^*(p) = \begin{cases} a_2 & \text{if } p \leq \frac{1}{4}, \\ a_0 & \text{if } \frac{1}{4} \leq p \leq \frac{3}{4}, \\ a_1 & \text{if } p \geq \frac{3}{4}. \end{cases}$$

Under Bayesian persuasion and cheap talk, the sender's maximal payoff is given by the concave and quasi-concave envelopes of the indirect utility evaluated at the prior (see Figure 1a). Their difference captures the value of commitment.



(a) Bayesian persuasion vs cheap talk

(b) Partial commitment payoff for $\pi = 0.5$

FIGURE 1. Comparison of benchmarks and partial commitment

Now consider persuasion with partial commitment. The sender chooses a vector of posterior beliefs that satisfies Bayes plausibility with weights λ *and that admits no profitable* *undetectable deviation. The marginal λ restricts the set of feasible posterior beliefs. For instance, let $\pi = (\frac{1}{2}, \frac{1}{2})$ and $\lambda = (\frac{1}{2}, 0, \frac{1}{2})$. Then any feasible pair of posteriors must satisfy $p_{a_1}(\omega_1) = \frac{1}{2} + \epsilon$ and $p_{a_2}(\omega_1) = \frac{1}{2} - \epsilon$ for some $\epsilon \in [-\frac{1}{2}, \frac{1}{2}]$.*[9] *The sender's optimal strategy is to induce the posterior beliefs $p_{a_2}(\omega_1) = 0$ and $p_{a_1}(\omega_1) = 1$ which results in an expected payoff of $\frac{3}{2}$. This is strictly below the Bayesian persuasion payoff, which equals $\frac{5}{3}$.*

Figure 1b illustrates the sender's maximal equilibrium payoff as a function of the marginal $\lambda = \lambda(a_1)$, restricting attention to marginals with $\lambda(a_0) = 0$. In particular, when $\lambda = (\frac{1}{3}, 0, \frac{2}{3})$, the sender attains the Bayesian persuasion payoff. Convex combinations across marginals yield the concave envelope $\mathrm{Cav}_\lambda\, e_S^*(\pi, \lambda)$ (dotted line). We show that in the dynamic game, the sender can achieve all such convex combinations (shaded region).

3.2. **Dynamic game.** In this section, we show that any equilibrium payoff of the static persuasion model with partial commitment and any convex combination of such payoffs, can be sustained as a uniform equilibrium payoff. The characterization depends only on the invariant distribution of the transition matrix.

The next example illustrates the key idea behind the construction. In particular, it shows that a sender with state-independent payoffs can attain the Bayesian persuasion payoff in the dynamic game.

Example 1 (Continued)**.** *Consider Example 1 in a dynamic environment. Suppose the states evolve according to*

$$Q = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

This Markov chain is pseudo-renewal, with invariant distribution $\mu(\omega_1) = \mu(\omega_2) = \frac{1}{2}$. Fix $q = (\frac{1}{3}, 0, \frac{2}{3})$, the distribution over recommendations induced by the Bayesian persuasion solution for prior $\mu = (\frac{1}{2}, \frac{1}{2})$.

In the first stage, the sender induces the posterior beliefs $p_{a_1} = (1, 0)$ and $p_{a_2} = (\frac{1}{4}, \frac{3}{4})$. The resulting prior beliefs in the next stage are $q_{a_1} = p_{a_1}Q = (\frac{2}{3}, \frac{1}{3})$ and $q_{a_2} = p_{a_2}Q = (\frac{5}{12}, \frac{7}{12})$.

Since the chain is pseudo-renewal, we have $q_{a_i} = \alpha p_{a_i} + (1 - \alpha)\mu$ for some $\alpha \in [0, 1)$. Hence, after every recommendation, the sender can again induce the same vector of posterior beliefs, namely p_{a_1} and p_{a_2}. The evolution of the receiver's beliefs is shown in Figure 2.

At every stage, the unconditional distribution over recommendations remains $(\frac{1}{3}, \frac{2}{3})$, and the posterior beliefs remain $p_{a_1} = (1, 0)$ and $p_{a_2} = (\frac{1}{4}, \frac{3}{4})$. By obedience, the receiver follows the recommendation: he takes action a_1 under p_{a_1} and a_2 under p_{a_2}. Moreover, the receiver monitors the frequency of recommendations and enforces that each action is played according to its prescribed quota λ. Since any undetectable deviation preserves

[9]Since $\lambda(a_0) = 0$, the posterior belief p_{a_0} is never realized and is therefore irrelevant.

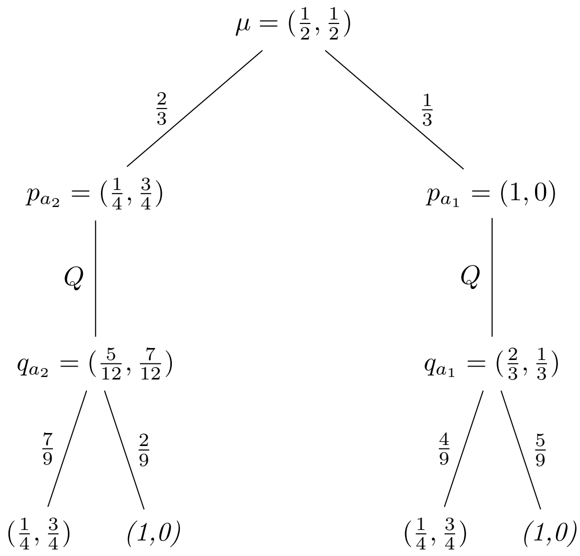


FIGURE 2. Evolution of the receiver's beliefs in Example 1

λ and corresponds to a feasible deviation in the static model, the sender has no profitable deviation. Thus, in equilibrium, the sender attains the Bayesian persuasion payoff $\text{Cav}\,\hat{u}_S(\mu) = \frac{5}{3}$.

We now state the main result. It establishes that any equilibrium payoff of the static persuasion model with partial commitment, with prior given by the invariant distribution, as well as any convex combination thereof, can be sustained as a uniform equilibrium payoff in the dynamic game. Let $\text{Co}_\lambda\big(\mathcal{E}(\mu,\lambda)\big)$ denote the set of convex combinations of equilibrium payoffs across distributions over messages.[10]

Theorem 1. *Any equilibrium payoff of the persuasion model with partial commitment and prior μ, as well as any convex combination of such payoffs across distributions over messages, is a uniform equilibrium payoff. Formally,*

$$\text{Co}_\lambda\big(\mathcal{E}(\mu,\lambda)\big) \subseteq \mathcal{U}.$$

Proof Sketch: We first show that any equilibrium payoff $e(\mu,\lambda)$ of the persuasion model with partial commitment is a uniform equilibrium payoff. We then show that any convex combination of such payoffs is also a uniform equilibrium payoff.

Fix $\lambda \in \Delta(A)$ and an equilibrium payoff $e(\mu,\lambda)$. Let $\{p_a\}_{a\in A}$ denote the equilibrium posterior beliefs, which satisfy Bayes plausibility under weights λ and the receiver's obedience condition.

We construct a strategy profile (σ^*,τ^*). The play is divided into consecutive blocks of length N. At the start of each block, players discard past play and restart playing a new block. Choose N such that $N\lambda(a)$ is an integer for all $a \in A$.

[10] $\text{Co}_\lambda\big(\mathcal{E}(\mu,\lambda)\big) = \{\sum_{i=1}^n \eta_i\, e(\mu,\lambda_i) \mid \lambda_i \in \Delta(A),\ \eta_i \ge 0,\ \sum_{i=1}^n \eta_i = 1\}$.

Receiver's strategy: At each stage n, the sender makes a recommendation m_n. If the recommended action has not yet reached its quota within the block, the receiver follows it; otherwise, he replaces it with an action whose quota is not yet filled.[11]

To keep track of action frequencies, let

$$C_n(a) := \sum_{k=1}^{n} \mathbf{1}_{\{a_k=a\}}$$

denote the number of times action a has been taken up to period n.

According to τ^*, if $C_{n-1}(m_n) < N\lambda(m_n)$, then the receiver plays the recommended action, i.e., $a_n = m_n$. Otherwise, he chooses an action $a \neq m_n$ such that $C_{n-1}(a) < N\lambda(a)$. For example, the action can be chosen randomly with probability proportional to the remaining quota $N\lambda(a) - C_{n-1}(a)$. Hence, in each block, action a is taken exactly $N\lambda(a)$ times.

Sender's strategy: The sender induces the same posterior beliefs $\{p_a\}_{a\in A}$ at every stage, so that the expected joint distribution over states and recommendations coincides with the equilibrium outcome of the persuasion model with partial commitment.

For each $a \in A$, let $q_a = p_a Q$ denote the next-period belief following recommendation a. Under the pseudo-renewal Markov chain, we have

$$q_a = \alpha p_a + (1-\alpha)\sum_{\tilde{a}\in A} \lambda(\tilde{a}) p_{\tilde{a}} = \alpha p_a + (1-\alpha)\mu,$$

so q_a lies on the segment joining p_a and μ.

Using this decomposition, the sender chooses the current recommendation as a function of the current state and the previous recommendation:

$$\sigma_n^*(m_n = a \mid \omega_n = \omega, m_{n-1} = a') := \frac{p_a(\omega)}{q_{a'}(\omega)}\Big((1-\alpha)\lambda(a) + \alpha \mathbf{1}_{\{a=a'\}}\Big).$$

The evolution of beliefs induced by the strategy is illustrated in Figure 3.

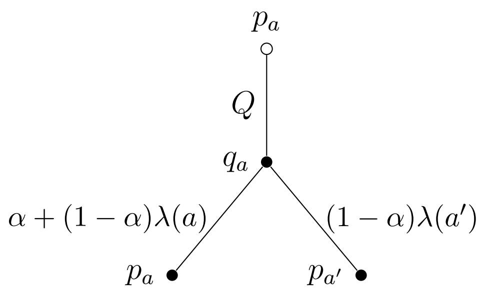


FIGURE 3. Recursive decomposition of receiver's beliefs under σ^*

[11]The receiver's strategy is based on statistical tests used in repeated games with imperfect monitoring (Lehrer, 1992a; Jackson and Sonnenschein, 2007; Renault, Solan, and Vieille, 2013; Escobar and Toikka, 2013; Frankel, 2016; Hörner, Mu, and Vieille, 2017). Deviations that alter the target distribution are detectable through these tests.

The proof proceeds in two steps. First, we show that the payoff induced by (σ^*, τ^*) can be made arbitrarily close to $e(\mu, \lambda)$ for sufficiently patient players. Second, we show that (σ^*, τ^*) is an approximate equilibrium for sufficiently patient players.[12]

Under (σ^*, τ^*), the expected joint distribution over states and recommendations at any stage coincides with the static distribution, i.e.,

$$\mathbb{E}_{\sigma^*,\tau^*}\left[\mathbf{1}_{\{\omega_n=\omega, m_n=a\}}\right] = \lambda(a)p_a(\omega).$$

Hence, for a large block N, recommendations and actions coincide with high probability in most stages. Therefore, by choosing N large and players sufficiently patient, the induced payoff can be made arbitrarily close to $e(\mu, \lambda)$.

Now, we show that this strategy profile is an approximate equilibrium when players are sufficiently patient.

First, consider the receiver's deviations. Since the sender does not observe the receiver's actions, the receiver's current choice does not affect future play. Hence, a myopic best response at each stage is optimal. Under τ^*, for a sufficiently large block N, the receiver follows the recommendation whenever the quota is not exhausted, which occurs in most stages. Therefore, τ^* coincides with a myopic best response in all but a small fraction of stages, so any gain from deviation is bounded by ϵ for sufficiently patient players.

Next, consider the sender's deviations. The receiver's strategy enforces that the frequency of action a in each block equals $N\lambda(a)$. Hence, any sender deviation $\tilde{\sigma}$ induces, within a block, a collection of conditional distributions over states $\{\tilde{p}_a\}_{a\in A}$, where

$$\tilde{p}_a(\omega) := \frac{1}{N\lambda(a)}\mathbb{E}_{\tilde{\sigma},\tau^*}\left[\sum_{n=1}^{N} \mathbf{1}_{\{\omega_n=\omega,\, a_n=a\}}\right].$$

These distributions satisfy Bayes plausibility with prior μ and weights λ, i.e., $\mu = \sum_a \lambda(a)\tilde{p}_a$.

This corresponds to a feasible deviation in the static model with marginal λ. By optimality of $\{p_a\}_{a\in A}$, such deviations cannot yield a higher payoff. Moreover, for sufficiently large N and sufficiently patient players, the sender's discounted payoff from any deviation is arbitrarily close to the payoff induced by the corresponding static deviation. Hence, no profitable deviation exists. Therefore, $e(\mu, \lambda) \in \mathcal{U}$.

We now extend the previous block construction to obtain convex combinations of equilibrium payoffs.[13] Let $\lambda = \sum_{i=1}^{n} \eta_i\lambda_i$ with $\eta_i \geq 0$ and $\sum_i \eta_i = 1$. The players still use a block strategy. However, each block of length N is further divided into sub-blocks of

[12] The strategy profile (σ^*, τ^*) also constitutes a uniform equilibrium when the sender observes the receiver's actions. This is related to Proposition 2 in Renault, Solan, and Vieille (2013), which shows that the set of Nash equilibrium payoffs without feedback is a subset of that with feedback.

[13] Block strategies are commonly used in dynamic games with imperfect monitoring; see, for example, Lehrer (1990, 1992b); Fudenberg and Levine (1991); Tomala (1999); Deb, González-Díaz, and Renault (2016).

lengths $\{N\eta_i\}_{i\in I}$, and in sub-block i the players play according to the strategy profile associated with λ_i.

Using the same argument as before, each sub-block yields an average payoff arbitrarily close to $e(\mu, \lambda_i)$ for sufficiently large N and sufficiently patient players. Since sub-block i occupies a fraction η_i of each block, the overall payoff is arbitrarily close to $\sum_{i=1}^{n} \eta_i e(\mu, \lambda_i)$. Moreover, because the strategy restarts at the beginning of each sub-block, deviations can be evaluated separately within each sub-block, and the same incentive argument applies.

3.3. **On the Converse and the Role of Dynamic Deviations.** Theorem 1 establishes that all payoff vectors that can be obtained as convex combinations of equilibrium payoffs of the persuasion model with partial commitment are uniform equilibrium payoffs. A natural question is whether the converse inclusion also holds.

A direct extension of the static benchmark would suggest that any outcome that admits a profitable deviation in the set $\Sigma(\mu, \lambda)$ of Bayes-plausible posterior vectors, holding fixed the long-run frequency of recommendations, should fail to be an equilibrium. However, in the dynamic environment, not all such deviations are feasible. Moreover, the receiver may condition on richer statistical tests, such as frequencies of pairs or longer strings of recommendations.

The following example shows that some uniform equilibrium outcomes lie outside the convex hull of equilibrium outcomes of the static persuasion model with partial commitment.

Example 2. *Consider binary states $\Omega = \{\omega_1, \omega_2\}$ and let $A = \{a_1, a_2, a_3, a_4\}$. The invariant distribution is $\mu = (\frac{1}{2}, \frac{1}{2})$ and is induced by the Markov chain*

$$Q = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Let the posteriors be $p_{a_1} = (0, 1)$, $p_{a_2} = (\frac{1}{3}, \frac{2}{3})$, $p_{a_3} = (\frac{2}{3}, \frac{1}{3})$, and $p_{a_4} = (1, 0)$, with $\lambda = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. These satisfy Bayes plausibility: $\sum_a \lambda(a) p_a = (\frac{1}{2}, \frac{1}{2})$. The belief dynamics are recursive: $p_{a_1} Q = p_{a_2}$ and $p_{a_4} Q = p_{a_3}$.

The sender's and receiver's payoffs are given by

	a_1	a_2	a_3	a_4
ω_1	$(-1, 8)$	$(3, 7)$	$(-2, 3)$	$(1, 0)$
ω_2	$(1, 0)$	$(-2, 3)$	$(3, 7)$	$(-1, 8)$

The receiver's optimal action is a_1 for $p \in [0, \frac{1}{4}]$, a_2 for $p \in (\frac{1}{4}, \frac{1}{2}]$, a_3 for $p \in [\frac{1}{2}, \frac{3}{4})$, and a_4 for $p \in [\frac{3}{4}, 1]$, where $p = \mathbb{P}(\omega_2)$ (with ties broken in favor of the sender).

In every odd period, the sender recommends a_1 or a_4, each with probability $1/2$. In every even period, she recommends a_2 after a_1 and a_3 after a_4. The receiver follows the recommendation as long as only the pairs (a_1, a_2) and (a_4, a_3) occur; if (a_1, a_3) or (a_4, a_2)

is observed, he ignores all future recommendations and plays his optimal action at the prior forever.

Along the prescribed path, the sender is truthful in odd periods. In even periods, however, she prefers to swap the continuation rule: at $p_{a_2} = (\frac{1}{3}, \frac{2}{3})$ *she prefers* a_3 *to* a_2*, while at* $p_{a_3} = (\frac{2}{3}, \frac{1}{3})$ *she prefers* a_2 *to* a_3*. Thus, she has a profitable deviation in the static persuasion model with fixed marginal* λ*, namely to send* a_3 *after* a_1 *and* a_2 *after* a_4*, which preserves the one-period distribution.*

In the dynamic game, this deviation is not feasible, since it changes the continuation pattern $a_1 \to a_2$*,* $a_4 \to a_3$ *into* $a_1 \to a_3$*,* $a_4 \to a_2$*, thereby altering the distribution over adjacent recommendation pairs. It is therefore detected and triggers the babbling outcome.*[14]

Hence, the induced payoff $\frac{1}{3}$ *is a uniform equilibrium payoff for the sender. We now show that it does not belong to the convex hull of sender's equilibrium payoffs of the persuasion model with partial commitment. By Proposition 2, any such payoff arises from a Bayes-plausible splitting of the prior and can be written as a convex combination of indirect utilities at posterior beliefs. The minimum of the sender's indirect utility is attained at the prior* π*, yielding* $\frac{1}{2}$*. Therefore, every payoff in the static benchmark is at least* $\frac{1}{2}$*, and thus* $u \in \mathcal{U}$ *but* $u \notin \mathrm{Co}_\lambda\big(\mathcal{E}(\mu, \lambda)\big)$*.*

This example shows that dynamically feasible deviations may be strictly smaller than those that preserve quotas. We therefore focus on a class of deviations that are always feasible and undetectable, yielding a necessary condition.

Consider any outcome $\nu \in \Delta(\Omega \times A)$ with prior $\pi \in \Delta(\Omega)$. Let

$$\mathcal{C}(\pi) := \left\{ c \in \Delta(\Omega \times \Omega) : \sum_{\omega'} c(\omega, \omega') = \pi(\omega) \text{ and } \sum_{\omega} c(\omega, \omega') = \pi(\omega') \right\}$$

denote the set of copulas with marginals π.

Define the set of state permutations

$$\mathcal{M}(\nu) := \left\{ \tilde{\nu} \in \Delta(\Omega \times A) : \tilde{\nu}(\omega, a) = \sum_{\omega'} c(\omega, \omega')\, \nu(\omega', a) \text{ for some } c \in \mathcal{C}(\pi) \right\}.$$

Definition 3. An outcome $\nu \in \Delta(\Omega \times A)$ with marginal π over states is **robust to state permutations** if

$$\mathbb{E}_\nu[u_S] \geq \mathbb{E}_{\tilde{\nu}}[u_S] \quad \text{for every } \tilde{\nu} \in \mathcal{M}(\nu).$$

State permutations form a subset of deviations that preserve the marginal distribution over messages. They correspond to relabeling states via a permutation and applying the equilibrium strategy to the resulting fictitious process. As shown in Renault, Solan, and

[14] A deviation that swaps the odd-period recommendations a_1 and a_4 while preserving the continuation pattern is feasible but not profitable. Such a deviation yields an average payoff of $\frac{-1+\frac{4}{3}}{2} = \frac{1}{6}$, which is strictly below the equilibrium payoff $\frac{1}{3}$.

Vieille (2013, Lemma 4), such deviations preserve the law of the state process and are therefore undetectable.

Proposition 2. *If an outcome $\nu \in \Delta(\Omega \times A)$ is induced by a uniform equilibrium, then ν satisfies obedience and is robust to state permutations. Moreover, its marginal over states equals the invariant distribution μ.*

Proposition 2 is not sufficient in general, since $\mathcal{M}(\nu)$ may be strictly smaller than the set of deviations in $\Sigma(\pi, \lambda)$[15] An immediate implication is that every uniform equilibrium outcome is a Bayes correlated equilibrium with prior μ (Bergemann and Morris, 2019).

In Renault, Solan, and Vieille (2013), state permutations are both necessary and sufficient to characterize equilibrium outcomes. In our setting, however, the sender recommends actions rather than reporting states, so deviations that preserve message frequencies may still alter the temporal structure of recommendations and become detectable. Thus, while permutation deviations remain necessary, they are not sufficient to characterize equilibrium outcomes, as illustrated in Example 2.

Finally, we consider the special case where the states are drawn i.i.d. with distribution μ. In this case, the dynamic restrictions discussed above disappear, and the static partial-commitment benchmark becomes exact.

Proposition 3. *Suppose that the state process is i.i.d. with distribution μ. Then the set of uniform equilibrium payoffs coincides with the convex hull of static persuasion payoffs with partial commitment:*

$$\mathcal{U} = \mathrm{Co}_\lambda\big(\mathcal{E}(\mu, \lambda)\big).$$

The inclusion follows from Theorem 1. For the converse, each period induces a splitting of μ into posterior beliefs with weights λ. If this splitting does not belong to $\mathcal{E}(\mu, \lambda)$, then the sender has a profitable undetectable deviation preserving λ. Since states are drawn i.i.d. across periods, this deviation can be implemented independently across periods, yielding a contradiction. Hence, the induced payoff must belong to $\mathcal{E}(\mu, \lambda)$, and the overall payoff is a convex combination of such payoffs across recommendation distributions.

4. Discussion and Conclusion

We studied a dynamic sender-receiver game, where the sender gets no feedback on the receiver's action. We partially characterize the set of uniform equilibrium payoffs by relating the dynamic game into a static persuasion model with partial commitment. Crucially, dynamic interaction helps the sender achieve a higher payoff despite no feedback. In particular, if the sender's payoff is state independent, she is able to achieve the Bayesian persuasion payoff.

[15]To see this, suppose $\rho(a \mid \omega) > 0$ for all ω, a, so that each posterior p_a has full support. State permutations preserve support and therefore cannot generate degenerate posteriors, whereas such posteriors are feasible in $\Sigma(\pi, \lambda)$.

Many problems remain open in this setting. First, fully characterizing the set of all uniform equilibrium payoffs remains an open problem. Second, exploring scenarios in which the sender receives imperfect feedback on the receiver's actions presents another avenue for investigation. Third, it would be interesting to study the role of more complex statistical tests by the receiver, such as tests based on higher-order frequencies or longer histories of messages (see Hörner, Mu, and Vieille, 2017), and how these affect the set of equilibrium outcomes. Additionally, extending the analysis to more general dynamic environments represents a promising direction.

References

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

Best, James and Daniel Quigley. 2024. "Persuasion for the long run." *Journal of Political Economy* 132 (5):1740–1791.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6):1431–1451. URL http://www.jstor.org/stable/1913390.

Deb, Joyee, Julio González-Díaz, and Jérôme Renault. 2016. "Uniform folk theorems in repeated anonymous random matching games." *Games and Economic Behavior* 100:1–23.

Escobar, Juan F and Juuso Toikka. 2013. "Efficiency in games with Markovian private information." *Econometrica* 81 (5):1887–1934.

Frankel, Alexander. 2016. "Discounted quotas." *Journal of Economic Theory* 166:396–444.

Fudenberg, Drew and David K Levine. 1991. "An approximate folk theorem with imperfect private information." *Journal of Economic Theory* 54 (1):26–47.

Green, Jerry R and Nancy L Stokey. 2007. "A two-person game of information transmission." *Journal of economic theory* 135 (1):90–104.

Hörner, Johannes, Xiaosheng Mu, and Nicolas Vieille. 2017. "Keeping your story straight: Truthtelling and liespotting." *Unpublished paper, Department of Economics, Harvard University.[609]* .

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

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review* 101 (6):2590–2615. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.6.2590.

Kuvalekar, Aditya, Elliot Lipnowski, and Joao Ramos. 2022. "Goodwill in communication." *Journal of Economic Theory* 203:105467.

Le Treust, Maël and Tristan Tomala. 2019. "Persuasion with limited communication capacity." *Journal of Economic Theory* 184:104940.

Lehrer, E. 1990. "Nash equilibria of n-player repeated games with semi-standard information." *International Journal of Game Theory* 19 (2):191–217. URL https://doi.org/10.1007/BF01761076.

Lehrer, Ehud. 1992a. "On the equilibrium payoffs set of two player repeated games with imperfect monitoring." *International Journal of Game Theory* 20 (3):211–226.

———. 1992b. "Two-player repeated games with nonobservable actions and observable payoffs." *Mathematics of Operations Research* 17 (1):200–224.

Lin, Xiao and Ce Liu. 2024. "Credible persuasion." *Journal of Political Economy* 132 (7):000–000.

Lipnowski, Elliot and Doron Ravid. 2020. "Cheap talk with transparent motives." *Econometrica* 88 (4):1631–1660.

Lipnowski, Elliot, Doron Ravid, and Denis Shishkin. 2022. "Persuasion via weak institutions." *Journal of Political Economy* 130 (10):2705–2730.

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

Min, Daehong. 2021. "Bayesian persuasion under partial commitment." *Economic Theory* 72 (3):743–764.

Renault, Jérôme, Eilon Solan, and Nicolas Vieille. 2017. "Optimal dynamic information provision." *Games and Economic Behavior* 104:329–349.

Renault, Jérôme, Eilon Solan, and Nicolas Vieille. 2013. "Dynamic sender-receiver games." *Journal of Economic Theory* 148 (2):502 – 534. URL http://www.sciencedirect.com/science/article/pii/S0022053113000161.

Tomala, Tristan. 1999. "Nash equilibria of repeated games with observable payoff vectors." *Games and Economic Behavior* 28 (2):310–324.

Appendix A. Omitted Proofs

A.1. **Proof of Proposition 1.** Let (ρ, κ) be an equilibrium of $\mathcal{P}(\pi, \lambda)$. We construct an equivalent equilibrium in which the sender directly recommends actions and the receiver obeys them.

Define

$$\tilde{\rho}(a \mid \omega) := \sum_{m \in M} \rho(m \mid \omega)\kappa(a \mid m) \qquad \text{for all } a \in A,\ \omega \in \Omega,$$

and let

$$\tilde{\kappa}(a \mid a) = 1 \qquad \text{for all } a \in A.$$

By construction, the induced joint distribution over states and actions is the same under (ρ, κ) and $(\tilde{\rho}, \tilde{\kappa})$. Hence, both players obtain the same expected payoff under the two strategy profiles. Moreover, the induced marginal distribution over recommendations is

$$\tilde{\lambda}(a) = \sum_{\omega \in \Omega} \pi(\omega)\tilde{\rho}(a \mid \omega) = \sum_{m \in M} \lambda(m)\kappa(a \mid m).$$

It remains to show that $(\tilde{\rho}, \tilde{\kappa})$ satisfies obedience and incentive compatibility.

First, consider the receiver. For each $a \in A$ such that $\tilde{\lambda}(a) > 0$,

$$\tilde{p}_a = \frac{1}{\tilde{\lambda}(a)} \sum_{m \in M} \lambda(m)\kappa(a \mid m)p_m,$$

where p_m denotes the posterior belief induced by message m under (ρ, κ). Thus, $\tilde{p}_a$ is a convex combination of the posteriors p_m for which $\kappa(a \mid m) > 0$. Since (ρ, κ) is an equilibrium, whenever $\kappa(a \mid m) > 0$, action a is a best response at p_m. By linearity, it follows that a is also a best response at $\tilde{p}_a$. Hence, obedience holds.

Next, consider the sender. Suppose, toward a contradiction, that there exists a profitable deviation under $(\tilde{\rho}, \tilde{\kappa})$, given by posterior beliefs $(\tilde{p}_a)_{a \in A}$ with weights $(\tilde{\lambda}(a))_{a \in A}$.

We construct a corresponding deviation in the original equilibrium by defining, for each message $m \in M$,

$$p_m := \sum_{a \in A} \kappa(a \mid m)\tilde{p}_a.$$

This deviation is feasible since

$$\sum_{m \in M} \lambda(m)p_m = \sum_{m \in M} \lambda(m) \sum_{a \in A} \kappa(a \mid m)\tilde{p}_a = \sum_{a \in A} \tilde{\lambda}(a)\tilde{p}_a = p.$$

Moreover, it yields the same expected payoff as the deviation in the direct policy, since

$$\sum_{m \in M} \lambda(m) \sum_{a \in A} \kappa(a \mid m) \sum_{\omega} \tilde{p}_a(\omega)u_S(\omega, a) = \sum_{a \in A} \tilde{\lambda}(a) \sum_{\omega} \tilde{p}_a(\omega)u_S(\omega, a).$$

Thus, this defines a profitable deviation in the original equilibrium, a contradiction. Therefore, $(\tilde{\rho}, \tilde{\kappa})$ is an equivalent equilibrium.

A.2. **Proof of Theorem 1.** First, we prove that any equilibrium payoff of the static persuasion model with partial commitment is a uniform equilibrium payoff. Then we show any convex combination of these payoffs with respect to the distributions over messages is also a uniform equilibrium payoff.

Fix $e(\mu,\lambda) \in \mathcal{E}(\mu,\lambda)$. Let (σ^*,τ^*) denote the strategy profile constructed in the proof sketch. The proof proceeds in two steps. First, we show that the payoff induced by (σ^*,τ^*) can be made arbitrarily close to $e(\mu,\lambda)$ for sufficiently patient players. Second, we show that (σ^*,τ^*) is an approximate equilibrium for sufficiently patient players.

Lemma 1. *For every $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ and $\delta_0 < 1$ such that, for every $N \geq N_0$ and $\delta \geq \delta_0$, the strategy profile (σ^*,τ^*) induces a payoff within ϵ of $e(\mu,\lambda)$.*

Proof. First, we show that at any stage n, the expected joint distribution over the state and the recommendation equals $\lambda(a)p_a(\omega)$.

$$\mathbb{E}_{\sigma^*}[\mathbf{1}_{\{\omega_n=\omega,m_n=a\}}] = \sum_{\tilde{a}} \sigma_n^*(m_n = a \mid m_{n-1} = \tilde{a}, \omega)\mathbb{P}_{\sigma^*,\tau^*}(m_{n-1} = \tilde{a}, \omega_n = \omega), \tag{A.1}$$

$$= \sum_{\tilde{a}} \frac{p_a(\omega)}{q_{\tilde{a}}(\omega)}\big((1-\alpha)\lambda(a) + \alpha\mathbf{1}_{\{\tilde{a}=a\}}\big)q_{\tilde{a}}(\omega)\lambda(\tilde{a}), \tag{A.2}$$

$$= \lambda(a)p_a(\omega). \tag{A.3}$$

where we used that $\mathbb{P}_{\sigma^*,\tau^*}(m_{n-1} = \tilde{a}, \omega_n = \omega) = \lambda(\tilde{a})q_{\tilde{a}}(\omega)$.

Define

$$\overline{p}_a(\omega) := \mathbb{E}_{\sigma^*,\tau^*}\Big[\frac{1}{N\lambda(a)}\sum_{t=1}^{N}\mathbf{1}_{\{\omega_t=\omega,a_t=a\}}\Big].$$

That is, $\lambda(a)\overline{p}_a(\omega)$ is the expected fraction of periods in a block in which the state is ω and the action is a.

For fixed N, since the strategy restarts every block, the discounted payoff converges to the average payoff within a block as $\delta \to 1$.

$$\lim_{\delta\to 1}\gamma_i^\delta(\sigma^*,\tau^*) = \sum_a \lambda(a)\sum_\omega \overline{p}_a(\omega)u_i(\omega,a) \quad \forall i \in \{S,R\}. \tag{A.4}$$

We will show that for sufficiently large block length, this payoff can be made arbitrarily close to the target payoff $e(\mu,\lambda)$.

First, observe that by a law of large numbers for Markov chains, if N is large enough, then with high probability the quota of each recommendation is exhausted only in a small fraction of periods. Hence, under (σ^*,τ^*), the expected fraction of periods in which recommendations and actions differ can be made arbitrarily small. It follows that $\overline{p}_a$ can be made arbitrarily close to p_a for every $a \in A$.

Using the fact that payoffs are bounded and the linearity of the expected payoff in RHS of (A.4), it follows that for every $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ and δ_0 such that for all

$N \geq N_0$ and $\delta \geq \delta_0$, it holds that

$$\|\gamma_i^\delta(\sigma^*, \tau^*) - e_i(\mu, \lambda)\| \leq \epsilon \quad \forall i \in \{S, R\}.$$

□

Lemma 2. *For every $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ and $\delta_0 \in (0,1)$ such that for all $N \geq N_0$ and $\delta \geq \delta_0$, the strategy profile (σ^*, τ^*) is an ϵ-equilibrium.*

Proof. Fix $\epsilon > 0$. We show that there exist N_0 and δ_0 such that for all $N \geq N_0$ and all $\delta \geq \delta_0$, no player can gain more than ϵ by deviating from (σ^*, τ^*).

Sender's deviations. Since the sender uses a block strategy σ^*, it suffices to evaluate deviations on a single block. Fix any deviation $\tilde{\sigma}$. Since under τ^* the receiver takes action a exactly $N\lambda(a)$ times in each block, define

$$\tilde{p}_a(\omega) := \frac{1}{N\lambda(a)} \mathbb{E}_{\tilde{\sigma},\tau^*}\left[\sum_{n=1}^{N} \mathbf{1}_{\{\omega_n=\omega, a_n=a\}}\right].$$

Then $(\tilde{p}_a)_{a\in A}$ satisfies Bayes plausibility with weights λ, that is,

$$\mu = \sum_{a\in A} \lambda(a)\tilde{p}_a.$$

Hence the collection $(\tilde{p}_a)_{a\in A}$ is feasible in $\mathcal{P}(\mu, \lambda)$.

For fixed N, as $\delta \to 1$, the sender's discounted payoff under $(\tilde{\sigma}, \tau^*)$ converges to the average payoff within a block. Therefore, for δ sufficiently close to 1,

$$\gamma_S^\delta(\tilde{\sigma}, \tau^*) \leq \sum_{a\in A} \lambda(a) \sum_{\omega\in\Omega} \tilde{p}_a(\omega) u_S(\omega, a) + \epsilon.$$

Since $(p_a)_{a\in A}$ is sender-optimal in $\mathcal{P}(\mu, \lambda)$, we have

$$\sum_{a\in A} \lambda(a) \sum_{\omega\in\Omega} \tilde{p}_a(\omega) u_S(\omega, a) \leq e_S(\mu, \lambda).$$

Combining the two inequalities gives

$$\gamma_S^\delta(\tilde{\sigma}, \tau^*) \leq e_S(\mu, \lambda) + \epsilon.$$

By lemma 1, for N large enough and δ sufficiently close to 1,

$$\gamma_S^\delta(\sigma^*, \tau^*) \geq e_S(\mu, \lambda) - \epsilon.$$

Therefore,

$$\gamma_S^\delta(\tilde{\sigma}, \tau^*) - \gamma_S^\delta(\sigma^*, \tau^*) \leq 2\epsilon.$$

After replacing ϵ by $\epsilon/2$, it follows that the sender cannot gain more than ϵ by deviating.

Receiver's deviations. Because the receiver's actions are unobservable, his choice at stage n does not affect the sender's future behavior. Hence, against σ^*, a myopic best response at every period is optimal. Let τ^{myp} denote the myopic best-response strategy: after every history, the receiver best responds in the current period to the belief p_t induced by the sender's recommendation.

It therefore suffices to show that deviating to τ^{myp} does not result in a gain greater than ϵ. By construction, τ^* follows recommendation a whenever the quota for a is not exhausted. By the law of large numbers for Markov chains, for N large enough the empirical frequency of recommendations is close to λ with high probability. Hence, τ^* coincides with τ^{myp} in all but an ϵ-fraction of periods in each block.

Since stage payoffs are bounded, this implies that for N large enough and δ sufficiently close to 1,

$$\gamma_R^\delta(\sigma^*, \tau^{\text{myp}}) - \gamma_R^\delta(\sigma^*, \tau^*) \le \epsilon.$$

Therefore, the receiver cannot gain more than ϵ by deviating. □

Lemma 1 and Lemma 2 show that any $e \in \mathcal{E}(\mu, \lambda)$ is a uniform equilibrium payoff. We now show that players can obtain any convex combination of such payoffs.

Pick any $e \in \text{Co}_\lambda\big(\mathcal{E}(\mu, \lambda)\big)$. Since payoffs lie in a finite-dimensional space, e can be written as a finite convex combination. Thus, there exist a finite index set I, weights $(\eta_i)_{i\in I}$ with $\eta_i \ge 0$ and $\sum_{i\in I} \eta_i = 1$, and marginals $\lambda_i \in \Delta(A)$ such that

$$e = \sum_{i\in I} \eta_i e(\mu, \lambda_i), \qquad \text{with } e(\mu, \lambda_i) \in \mathcal{E}(\mu, \lambda_i) \text{ for each } i \in I.$$

The players use a block strategy. However, each block of length N is further subdivided into I sub-blocks.[16] Sub-block i starts at period $N\sum_{j=1}^{i-1} \eta_j + 1$ and ends at period $N\sum_{j=1}^{i} \eta_j$. The players play according to strategy profile (σ_i^*, τ_i^*) in sub-block i.

Using the same argument as in Lemma 1, for each sub-block i, taking N sufficiently large ensures that the payoff induced by (σ_i^*, τ_i^*) is arbitrarily close to $e(\mu, \lambda_i)$ for all δ sufficiently close to 1. Since sub-block i occupies a fraction η_i of each block, it follows that

$$\lim_{\delta\to 1} \gamma^\delta(\tilde\sigma, \tilde\tau) = \sum_{i\in I} \eta_i e(\mu, \lambda_i).$$

It remains to check incentives. Because the strategy restarts at the beginning of each sub-block, deviations in one sub-block do not affect continuation play in subsequent sub-blocks. Thus, any unilateral deviation decomposes into deviations within sub-blocks.

By Lemma 2, for every $i \in I$, there exists δ_0^i such that for every $\delta \ge \delta_0^i$, no player can gain more than $\eta_i\epsilon$ by deviating in sub-block i. Let $\delta_0 = \max_{i\in I} \delta_0^i$. Then for every $\delta \ge \delta_0$, the total discounted gain from deviating is at most $\sum_{i\in I} \eta_i \epsilon = \epsilon$.

[16] Choose N such that $N\eta_i\lambda_i(a)$ is an integer for all $i \in I$ and $a \in A$.

Using the convergence result, this implies that the difference in payoffs between the equilibrium strategy and any deviation is bounded above by 2ϵ. Therefore,

$$\sum_{i=1}^{I} \eta_i e(\mu, \lambda_i) \in \mathcal{U}.$$

A.3. **Proof of Proposition 2.** Fix $\epsilon > 0$, and let (σ^*, τ^*) be an ϵ-equilibrium of the δ-discounted game for all $\delta \geq \delta_0$. Let

$$\nu(\omega, a) := (1-\delta) \sum_{n=1}^{\infty} \delta^{n-1} \mathbb{E}_{\sigma^*,\tau^*}[\mathbf{1}_{\{\omega_n=\omega, a_n=a\}}]$$

denote the outcome induced by the strategy profile. Since the initial state is drawn from the invariant distribution μ, it follows that $\sum_a \nu(\omega, a) = \mu(\omega)$. Using this, denote

$$\rho_{\sigma^*,\tau^*}(a \mid \omega) := \frac{1}{\mu(\omega)} \sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} \mathbb{E}_{\sigma^*,\tau^*}[\mathbf{1}_{\{\omega_n=\omega, a_n=a\}}].$$

Consider the deviation τ^{myp} where the receiver plays a myopic best response to belief p_n at every history. Since continuation payoffs do not depend on the receiver's action, we have

(A.5)

$$\gamma_R(\sigma^*, \tau^{\text{myp}}) - \gamma_R(\sigma^*, \tau^*) = \mathbb{E}_{\sigma^*,\tau^*}\left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} \Big(\mathbb{E}_{p_n}[u_R(\cdot, a^*(p_n))] - \mathbb{E}_{p_n}[u_R(\cdot, a_n)]\Big)\right].$$

Since (σ^*, τ^*) is an ϵ-equilibrium, the left-hand side is at most ϵ.

Fix any $\eta > 0$, and let

$$R_\eta^n := \{\mathbb{E}_{p_n}[u_R(\cdot, a^*(p_n))] - \mathbb{E}_{p_n}[u_R(\cdot, a_n)] \geq \eta\}.$$

denote the event when the receiver's deviation in period n is at least η.

Then (A.9) implies

$$\epsilon \geq \mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} \Big(\mathbb{E}_{p_n}[u_R(\cdot, a^*(p_n))] - \mathbb{E}_{p_n}[u_R(\cdot, a_n)]\Big)\right] \geq \eta\, \mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} \mathbf{1}_{R_\eta^n}\right].$$

Hence,

$$\mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} \mathbf{1}_{R_\eta^n}\right] \leq \frac{\epsilon}{\eta}.$$

Therefore, for every $\eta > 0$, the discounted weight of periods in which the receiver's action is worse than a myopic best response by at least η vanishes as $\epsilon \to 0$.

In case of the sender, we show that there is an undetectable deviation σ', which needs to not be profitable more than ϵ. Using Lemma 4 in Renault, Solan, and Vieille (2013), one can construct a sequence of fictitious states (θ_n) that is statistically indistinguishable from the sequence (ω_n). In particular, (i) the law of $\{\theta_n\}_n$ is the same as the law of $\{\omega_n\}_n$, (ii) in each period n, the law of the pair (ω_n, θ_n) is given by distribution $c \in \mathcal{C}(\mu)$ and (iii) the conditional law of ω_n given $\theta_1, \ldots, \theta_n$ is given by $c(\cdot \mid \theta_n)$. Moreover, the sequence θ_n can be constructed using informationa available only in period n. Given such a sequence, in any stage n the sender will play according to $\sigma(\cdot \mid \theta_n)$ after replacing the realized state ω_n by the fictitious state θ_n. Formally, given history $(\omega_1, \theta_1, m_1, \ldots, \omega_n, \theta_n)$ consisting of realized and fictitious states and messages, the strategy $\sigma^{'}$ plays according to $\sigma(\theta, m_1, \ldots, \theta_n)$ that would have been played by σ had the realized states and messages been $\theta_1, m_1, \ldots, \theta_n$.

We now show that the sender's expected payoff under the strategy profile $(\sigma^{'}, \tau)$ is equal to $\sum_{\omega,\theta} c(\omega, \theta) \sum_a \sigma(a \mid \theta) u_S(\omega, a)$.

$$\mathbb{E}_{\sigma^{'},\tau^*}[u(\omega_n, a_n)] = \sum_{\omega_n} \sum_{\theta_1,\ldots,\theta_n} \sum_{a_1,\ldots,a_n} \mathbb{P}_{\sigma^{'},\tau}(\theta_1, a_1, \ldots., \omega_n, \theta_n, a_n) u_S(\omega_n, a_n) \tag{A.6}$$

$$= \sum_{\omega_n} \sum_{\theta_1,\ldots,\theta_n} \sum_{a_1,\ldots,a_n} \mathbb{P}_{\sigma^{'},\tau}(\omega_n \mid \theta_1, \ldots \theta_n) \times \mathbb{P}_{\sigma^{'},\tau}(\theta_1, a_1 \ldots \theta_n, a_n) u_S(\omega_n, a_n) \tag{A.7}$$

$$= \sum_{\omega_n} \sum_{\theta_n} \sum_{a_n} c(\omega_n \mid \theta_n) \times \mathbb{P}_{\sigma^{'},\tau}(\theta_n, a_n) u_S(\omega_n, a_n) \tag{A.8}$$

Since the process (θ_n) has the same law as (ω_n), and since σ' applies σ^* to the fictitious history, the discounted joint distribution over (θ_n, a_n) under (σ', τ^*) coincides with the discounted joint distribution of (ω_n, a_n) under (σ^*, τ^*). Hence,

$$(1-\delta) \sum_{n=1}^{\infty} \delta^{n-1} \mathbb{P}_{\sigma',\tau^*}(\theta_n = \theta, a_n = a) = \mu(\theta) \rho_{\sigma^*,\tau^*}(a \mid \theta).$$

Therefore,

$$\gamma_S^\delta(\sigma', \tau^*) = \sum_{\omega,\theta,a} c(\omega \mid \theta) \mu(\theta) \rho_{\sigma^*,\tau^*}(a \mid \theta) u_S(\omega, a).$$

So, for any outcome to be an equilibrium, there must be no profitable deviation in the set of deviations defined by $\mathcal{M}(\nu)$.

A.4. **Proof of Proposition 3.** The inclusion follows from Theorem 1. We prove the converse.

Fix $\epsilon > 0$, and let (σ, τ) be an ϵ-equilibrium of the δ-discounted game for all $\delta \geq \delta_0$.

Consider the deviation τ^{myp} where the receiver plays a myopic best response to belief p_n at every history. Since continuation payoffs do not depend on the receiver's action, we have

$$\gamma_R(\sigma, \tau^{\text{myp}}) - \gamma_R(\sigma, \tau) = \mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} \Big(\mathbb{E}_{p_n}[u_R(\cdot, a^*(p_n))] - \mathbb{E}_{p_n}[u_R(\cdot, a_n)] \Big)\right]. \tag{A.9}$$

Since (σ, τ) is an ϵ-equilibrium, the left-hand side is at most ϵ.

Fix any $\eta > 0$, and let

$$R_\eta^n := \{\mathbb{E}_{p_n}[u_R(\cdot, a^*(p_n))] - \mathbb{E}_{p_n}[u_R(\cdot, a_n)] \geq \eta\}.$$

denote the event when the receiver's deviation in period n is at least η.

Then (A.9) implies

$$\epsilon \geq \mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty}(1-\delta)\delta^{n-1}\Big(\mathbb{E}_{p_n}[u_R(\cdot,a^*(p_n))]-\mathbb{E}_{p_n}[u_R(\cdot,a_n)]\Big)\right] \geq \eta\,\mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty}(1-\delta)\delta^{n-1}\mathbf{1}_{R^n_\eta}\right].$$

Hence,

$$\mathbb{E}_{\sigma,\tau}\left[\sum_{n=1}^{\infty}(1-\delta)\delta^{n-1}\mathbf{1}_{R^n_\eta}\right] \leq \frac{\epsilon}{\eta}.$$

Therefore, for every $\eta > 0$, the discounted weight of periods in which the receiver's action is worse than a myopic best response by at least η vanishes as $\epsilon \to 0$.

Consider the sender's deviations. At any history $h_n = (m_1, \ldots, m_n)$, the sender's strategy induces a marginal distribution $\lambda_{h_n} \in \Delta(A)$ over action recommendations. Since the state process is i.i.d., the receiver's belief before observing the message is always equal to the prior μ.

Consider the deviation σ', where at any history h_n, if there exists a profitable deviation in $\Sigma(\mu, \lambda_{h_n})$, the sender switches to one that achieves the highest profit, and otherwise follows σ. Such a deviation is undetectable because it keeps the distribution over the message process unchanged. Also, from the i.i.d assumption, it does not affect continuation payoffs since future states are independent of the current one.

Like before, fix any $\eta > 0$ and let S^n_η be the event that in period n there exists a feasible one-shot deviation in $\Sigma(\mu, \lambda_n)$ that improves the sender's current payoff by at least η.

We have

$$\gamma_S(\sigma', \tau) - \gamma_S(\sigma, \tau) \geq \eta \cdot \mathbb{E}_{\sigma,\tau}\Big[\sum_{n=1}^{\infty}(1-\delta)\delta^{n-1}\mathbf{1}_{S^n_\eta}\Big]$$

Since (σ, τ) is an ϵ-equilibrium, it follows that

$$\mathbb{E}_{\sigma,\tau}\Big[\sum_{n=1}^{\infty}(1-\delta)\delta^{n-1}\mathbf{1}_{S^n_\eta}\Big] \leq \frac{\epsilon}{\eta}.$$

Since this holds for every $\eta > 0$, the discounted weight of stages at which the sender has a deviation gain bounded away from zero vanishes as $\epsilon \to 0$. It follows that the sender's static incentive constraint holds in the limit.

Thus, for any $\eta > 0$, the discounted weight of periods in which either the sender's or the receiver's equilibrium condition is violated by at least η is at most $2\epsilon/\eta$. Hence, up to a set of periods with vanishing discounted weight, the induced stage outcome in each period satisfies the equilibrium conditions of the static persuasion model with prior μ and marginal λ_{h_n}. Therefore, it follows that any limit point of the induced discounted payoffs belongs to

$$\mathrm{Co}_\lambda\big(\mathcal{E}(\mu, \lambda)\big).$$