DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 24-14

Receiver Inattention and Persuading to be Persuaded

Mengqi Zhang
University of Colorado Boulder

November 3, 2024
Revised November 13, 2024

Department of Economics

University of Colorado Boulder
Boulder, Colorado 80309

Receiver Inattention and Persuading to Be Persuaded

Mengqi Zhang*

November 13, 2024

Please always check the Latest Version

Abstract

Frictions in a Bayesian persuasion game, such as the receiver's rational inattention, can constrain the feasible information structures beyond Bayes' plausibility. In a conventional persuasion scenario with a binary state and binary action, we examine the properties of the inattention constraint under which the sender is likely to benefit from extending the persuasion game. These properties transform the sender's persuasion problem into an intertemporal one, where her strategy not only determines the current chance to succeed but also the receiver's prior belief in the next persuasion attempt, if necessary. In contrast to the optimal static persuasion strategy, the intertemporal approach may lead the sender to adopt a "piecemeal" information disclosure strategy, where she sacrifices the chance of immediate success to ensure that the receiver can be persuaded in subsequent attempts should her current attempt fail. While extending the persuasion game can improve overall persuasiveness beyond the static efficiency level, frictional constraints continue to define the efficiency limits of this sequential strategy. Friction-free efficiency remains unattainable, even if the limit on the opportunities of persuade is non-binding.

*Department of Economics, University of Colorado Boulder. Email: mengqi.zhang@colorado.edu. I am grateful to Jin-Hyuk Kim, Piotr Dworczak, Daniel Rappoport, Ernesto Rivera Mora, Yangwei Song, David Dillenberger, Luciano Pomato, Shaowei Ke, Romans Pancs, and the audience at the 11th Hong Kong Economic Association Biennial Conference at the University of Hong Kong for their insightful comments and advice. I bear complete responsibility for all the errors present in the paper.

"The purpose of an elevator pitch is to describe a situation or solution so compelling that the person you're with wants to hear more even after the elevator ride is over."

—-*Seth Godin*[1]

1 Introduction

Bob's decision, based on his belief about the true state, can affect Alice's payoff. By designing the information structure—a rule that maps possible states to different signals—Alice can shape Bob's belief, thereby changing his decision to maximize her payoff. In the canonical Bayesian persuasion model, Alice can commit to the signal revelation rule, and Bob processes the signal according to the Bayes' rule, making Bayes plausibility the only requirement for an information structure to be feasible.

This frictionless framework, despite its elegance, may oversimplify the real-world complexities. Alice and Bob may meet in an elevator. Since Alice's opportunity to present her idea is constrained to a few minutes during the elevator ride, she may not be able to deliver her best pitch. In a Bayesian persuasion game, frictions like the receiver's inattention may make a Bayes-plausible information structure infeasible. If Alice chooses to pitch as if she were in a frictionless scenario, she may be unable to finish her pitch, or Bob may only process the signal partially (Bloedel and Segal, 2018), making persuasion less effective. How should a sender beat the receiver's inattention in a Bayesian persuasion game?

Che et al. (2023) examined friction in a sequential Bayesian persuasion setting, observing that belief-changing signals may not arrive with every otherwise successful attempt. This forces the sender to persuade sequentially and keep the receiver engaged over time. However, this observation may not represent friction itself but rather the sender's strategy to overcome it. If Alice struggles to convince Bob during an elevator ride, she may adopt a strategy that provides belief-changing information during the elevator ride while allowing room to continue persuading him if the initial attempt is unsuccessful—even at the expense of a belief-changing signal arriving on the first attempt.

Persistence is key to success in challenging persuasion tasks. Steve Jobs spent every weekend for five months persuading former PepsiCo CEO John Sculley to join Apple. Given the informal nature of these meetings, nearly each attempt was an "elevator pitch," with only limited attention from Sculley. Since persuasion rarely succeeds on the first attempt, it is important to make the receiver persuadable on the following attempts, just as Seth Godin explains with the elevator pitch. In contrast to most previous relevant studies where the game duration is exogenously determined, we assume that the sender can strategically choose the information structure to extend the persuasion game. In this framework, where the duration of the game is endogenous, we investigate what properties of the frictional constraint will motivate the sender to extend the game for higher effectiveness in persuasion.

The significance of whether sequential persuasion enhances persuasiveness hinges on another key question: how much does a sequential strategy—especially the first attempt—differ from the strategy that would optimize a one-shot persuasion game? If the optimal opening pitch in a sequential approach aligns with the best strategy for concluding persuasion in a single attempt, Alice need not worry about whether sequential persuasion improves her overall effectiveness. She can simply focus on persuading as hard as possible within the current constraint; if she fails, she waits for another opportunity from Bob to try again. However, this may not be the case. To secure a second chance in the case that her first attempt fails, Alice might need to adjust her opening pitch strategically to ensure that Bob's patience is not exhausted by her first failure. In such cases, overlooking the benefits of sequential persuasion could reduce Alice's overall effectiveness—or even cost her a second chance to persuade.

The intuition behind designing an optimal static versus sequential persuasion strategy generally differs, particularly in the sender's first attempt. In a sequential design, the first attempt serves as an opening, whereas in a static strategy, it is the sender's final opportunity. In some frictionless, static persuasion problems, maximizing the chance of success involves making the "bad signal" as bad as possible. As shown by Kamenica and Gentzkow (2011), if a good state does not produce bad signals, the sender can design the bad state to send a signal with maximum probability while the signal remains "good," which recommends the receiver to "act." This design maximizes the probability of success within the single attempt. However, if this attempt fails, the bad state is identified, leaving no room for further persuasion. This logic holds in a frictional setting where the receiver is rationally inattentive. If Alice aggressively persuades in her first attempt, aiming for a one-shot success, the failure would lead Bob to believe that any future persuasion attempt is unlikely or impossible to change his mind. Consequently, if Alice anticipates the difficulty of immediate success and values the chance to persuade again, she might adopt a more conservative approach, withholding some information in her initial pitch to preserve a backup opportunity.

To keep the receiver's patience and make him open to further persuasion, it is common in practice to keep an ace up the sender's sleeve. For example, a job candidate prepares multiple versions of her pitch. While each version provides a complete analysis of the same topic, the time required to finish the pitch—such as 1-minute, 5-minute, or 10-minute—determines the level of analytical detail. The candidate discloses more details one at a time, until she either convinces the search committee or is interrupted. This "piecemeal" information disclosure is proven to be effective when the interviewer has limited time and attention. It endogenously explains the observation by Che et al. (2023): each attempt offers incremental insights that rarely result in a "jump" over the cutoff in the receiver's belief but still serve to persuade gradually while preserving the receiver's patience. This approach allows for multiple attempts and technically expands the receiver's participation constraint, thereby increasing the overall chance of success.

Our study is established on the canonical persuasion problem with binary states and binary actions. We modify this game by introducing a frictional constraint that restricts information structures beyond Bayes plausibility. Additionally, the sender is allowed to propose and conduct multiple experiments, as long as they do not exceed the attempt limit and the receiver remains attentive to them. This endogenous sequential framework turns the persuasion game into an intertemporal problem for the sender. By designing the information structure, the sender not only determines her success rate in the current attempt but also shapes the receiver's prior belief, should further persuasion attempts be necessary.

Using this model, we investigate the conditions under which a two-stage persuasion strategy is more effective than a static approach (Proposition 1). This benchmark shows that an endogenous sequential framework is necessary only when the frictional constraint renders some Bayes-plausible information structures infeasible. Otherwise, a sequential strategy never outperforms an optimal static one. Naturally, the receiver may become disappointed after a failed attempt, reducing the feasible set of information structures for future persuasion efforts (Proposition 2). If the receiver's motivation is sensitive to prior failures, the sender must adopt a conservative persuasion strategy in the initial attempt (Proposition 4); otherwise, she risks losing effectiveness or even the chance to persuade again (Proposition 3). Generalizing this finding, we examine conditions under which the sender benefits from extending the persuasion path beyond two stages (Proposition 5), where she chooses a "piecemeal" information disclosure strategy (Proposition 6). To explore to what extent a sequential strategy can beat frictional constraints due to the receiver's inattention, we allow for a sufficiently large attempt limit. While prolonging the persuasion path can enhance effectiveness, we find that the efficiency boundary remains determined by the frictional constraint (Proposition 7).

The endogenous dynamic nature of this study distinguishes it from related research, particularly those that also examine multiple signals for a single non-stochastic state. Here, the sender fully controls the length of the persuasion path,[2] with the option to make it either purely static or sequential with any duration. While signal realizations play a crucial role in prompting the sender to end the game, in this study's specific problem, only one of the two possible signals in each persuasion attempt triggers termination. Therefore, as long as the game is active, the persuasion process strictly follows the path set by the sender. It is the sender—not the game environment—who determines how the receiver's prior belief evolves beyond the initial attempt. When the prior belief is considered a "state variable" at each period, the sender's decision makes the problem intertemporal. Unlike typical intertemporal problems, here the sender also chooses the optimal stopping point for the persuasion process when the attempt limit is sufficiently large to be non-binding.

[2]It is important to distinguish the length of the persuasion path from the length of the persuasion. The former determines when the sender finishes the persuasion or gives up regardless of the outcome, which is purely the sender's decision. Besides the sender's decision, the latter is also determined by the experiment, where the good signal terminates the subsequent persuasion effort immediately.

This special intertemporal framework, incorporating optimal stopping and within a persuasion problem, provides theoretical insights into decision-making in information design. By allowing the sender's responses to frictional constraints to manifest, it explains the sender's conservative persuasion behaviors, such as "piecemeal information disclosure," which might otherwise be mistaken for friction itself. This insight not only may shift researchers' perspectives on "slow persuasion," but also has potential policy implications, offering policymakers criteria to monitor or regulate persuasion practices to enhance social welfare. Beyond explaining the sender's decision, a more important contribution of this framework is its ability to connect the receiver's inattention and other frictions—often unobservable—to the observable decisions. With this theoretical foundation, calibrating parameters in the information cost as a function of mutual information becomes possible. It may pave the way for future studies to better understand decision-makers' inattention and its impacts across various fields.

To study the endogenous sequential Bayesian persuasion game, we develop a system that visualizes the feasible information structures as constraints, along with indifference curves representing success probabilities based on prior beliefs (Figure 2). This system enables direct comparisons of the persuasiveness of strategies with varying attempts and provides an intuitive way—through analyzing boundary shapes and shifts—to characterize constraint properties that influence sender decisions. Based on these intuitive characterizations, the system also offers algorithms to help the sender find optimal information structures for constructing an effective persuasion path under different frictional constraints. With its flexible approach, this system can be readily applied to a broad range of scenarios, which may be beyond the scope of this study but can be formalized with a similar theoretical framework, providing tractable analysis or intuitive interpretation.

1.1 Related Literature

Kamenica and Gentzkow (2011) established the framework for static persuasion where the sender has full commitment as she designs the experiment ex ante in the context of symmetrically incomplete information. Recent studies have developed this framework from static into dynamic. Among these studies, Che et al.(2023), Su et al.(2022), Ni et al. (2023) are most relevant studies to our research. All three studies examined the scenario where the sender designs experiments to induce the ideal prior belief for the subsequent persuasion attempts. Information costs play an important role in Che et al. (2023) but the dynamism of their game is introduced by the uncertain arrival of the effective signal. Su et al.(2022) discussed how sequential structure in the persuasion game may expand the sender's constrained signal space but the sequential structure is predetermined and the constrained is set by experiments. Since it is not the sender's choice to induce a subsequent persuasion attempt, the trade-off between immediate success and opportunity for subsequent attempt in persuasion was not fully characterized in this study.

In a more general framework, Ni et al. (2023) established the mathematical concept that the "frictional" constraint beyond Bayes plausibility can be expanded through sequential persuasion, where the sender chooses the sequence and may terminate it at a certain point by choosing trivial experiments afterwards. They mainly focus on the possibility and the limit of an "endogenous" sequential effort surpassing the constraint. However, whether the sender favors a sequential approach over the static one in persuasion is largely determined by the properties of this frictional constraint in the specific problem, which also shape the sender's design of the persuasion strategy. But they remain largely unexplored.[3] To address this gap, our paper situates the problem in a concrete economic context, in which we can characterize the shape of the frictional constraint and how it evolves over the course of persuasion. This approach not only allows us to explore and discuss the conditions and mechanisms under which the sender might strategically prolong the persuasion process, but it also provides an intuitive microeconomic foundation for the sender's behavior highlighted in related research, such as Che et al. (2023).

Rather than assuming a static state space with dynamic signal, some studies assume a stochastic change in state or state space over time (Ely, 2017; Senkov, 2022), sometimes in

[3]Furthermore, the feasible set in their research, due to being very general, does not exclude specific cases that inherently favor lengthy sequential persuasion, allowing the sender to maintain some ad-hoc motivation to extend the game.

accordance with Markov Chain (Renault et al., 2017; Ashkenazi-Golan et al., 2022; Lehrer and Shaiderman, 2022). Some other literature emphasizes the sender's choice to prolong the persuasion process either because the duration of persuasion improves the final decision (Bizzotto et al. 2020; Senkov, 2022), the duration of persuasion itself is profitable (Ely and Szydlowski, 2020; Orlov et al. 2020), or the duration eliminates unfavorable receiver's types from the game (Honryo, 2018; Guo and Shmaya, 2018). In these works, the frictional constraint or the receiver's inattention does not play an essential role and the sender's intention to extend persuasion into dynamic or one with more sequences is not to improve the information structures' feasibility.

2 A Toy Example

To illustrate how a sequential persuasion strategy can improve persuasion effectiveness when information structures are subject to frictional constraints, we present a toy numerical example. In the elevator pitch example, assume Alice is now a startup founder. She is pitching her idea to Bob, a general partner at a venture fund. Alice's business idea has potential—it could be successful on the market, generating a 1 (billion dollar) payoff for the investor, or it could fail, resulting in 0 value. Although Alice understands her business idea, she is inexperienced and uncertain about its market viability. In this case, the information is symmetric between Bob and Alice at the beginning. They both believe the chance of success is 0.5. To launch her business, Alice requests a 0.6 (billion dollar) investment from Bob. Because the initial expected value of the business idea is less than the requested funds, Bob will decline the request if no further information is provided.

To change Bob's decision, Alice must introduce her idea in greater detail. Bob then evaluates the information provided and gives honest feedback based on his assessment. Alice can present her business idea in three ways: S, A, and B. Each of these ways is characterized by an information structure as follows, where both high-profit (h) and low-profit (l) business models can be recognized as a good (g) or bad (b) idea. Here, π denotes the conditional probability of these events.

	Information Structure		
	S	A	B
$\pi(g\|h)$	1	0.8	0.6
$\pi(b\|h)$	0	0.2	0.4
$\pi(g\|l)$	2/3	0.2	0.2
$\pi(b\|l)$	1/3	0.8	0.8

Table 1: Three Ways to Introduce the Business Idea

In the frictionless Bayesian persuasion game, S is the best information structure to persuade Bob. If the signal g appears, it increases Bob's belief (that the business idea works) to 0.6, prompting him to invest. When signal b appears, he rejects the request. Under this information structure, Alice has a 5/6 chance to get the requested investment.

However, the information structure S may be too complex to deliver within the elevator ride, possibly due to the technical detail required to ensure that a highly profitable idea is never mistaken for a bad one. Suppose that if Alice fails to complete her first pitch during the elevator ride, Bob loses interest due to the perception that she is unable to clearly introduce the idea. In this particular scenario, the information structure S is not feasible. Fortunately, the simpler information structures A and B meet the elevator pitch time constraint. If Alice successfully completes her initial pitch in the elevator with either information structure, she will gain a second opportunity to continue her pitching outside of the elevator if necessary, where the feasible information structures are still A and B.

In a single persuasion attempt, the information structure A outperforms B: it offers a 0.5 probability of generating the g signal, changing Bob's belief to 0.8, which leads to the approval of the investment. In contrast, the information structure B provides only a 0.4 chance of convincing Bob to invest with the signal g.

However, the information structure A loses its advantage when a subsequent persuasion chance is considered. If Alice starts with A and fails, Bob's belief drops to 0.2 upon receiving the signal b. This prior belief is too low for Alice to change Bob's decision with a second persuasion attempt, regardless of whether she uses A or B. Specifically, the signal g from A or B would only raise Bob's belief to 0.5 or 3/7, respectively—both below the 0.6 threshold needed to secure his investment. Thus, if Alice begins with A, her overall chance of success stays at 0.5, as the failure in the first attempt also eliminates any further opportunities to succeed.

By contrast, starting the persuasion with B preserves a second chance for Alice to succeed. If the signal b appears in Alice's first persuasion attempt, Bob's belief only falls to 1/3. With a subsequent attempt using A, the signal g will appear with a probability of 2/5, leading Bob's belief to surpass the 0.6 threshold. Overall, Alice will have a $0.4 + 0.6 \times 2/5 = 0.64$ chance of obtaining the fund from Bob, which is higher than 0.5 had Alice started the persuasion using A.

In a more challenging scenario, Bob is disappointed and loses patience when Alice fails to impress him in the elevator. Due to the disappointment, Bob only accepts a second pitch using the original information structure B. Even under this restriction, using B to persuade Bob again still results in a 1/3 chance of generating the signal g, which changes Bob's belief from 1/3 to 0.6 and advises him to approve the investment. Keep persuading Bob with the information structure B gives Alice an overall chance of $0.4 + 0.6 \times 1/3 = 0.6$ to convince Bob, which remains higher than 0.5. The following table summarizes the expected persuasion outcomes.

First Attempt		**Second Attempt**		**Overall Probability to succeed**
Strategy	**Probability to succeed**	**Strategy**	**Probability to succeed**	
A	0.5	B	0	0.5
A	0.5	A	0	0.5
B	0.4	A	2/5	0.64
B	0.4	B	1/3	0.6

Table 2: Expected Persuasion Outcome in A Two-Stage Framework

This toy example, despite its stylized feasible set, provides intuitive answers to the main questions of this research. When friction restricts the feasibility of information structures and renders the optimal frictionless option infeasible, a sequential persuasion design can outperform the optimal static strategy under the frictional constraint, as represented by A in this example. This remains valid even when a previous failure in persuasion makes the receiver less patient, thereby further narrowing the feasible set of information structures. Moreover, the sender may need to strategically choose an information structure that appears less effective within the static framework to leverage the effectiveness of the sequential persuasion strategy. In this example, while the information structure A offers a promising chance of success in the initial attempt, it is overly aggressive considering that there is a potential second chance; if A fails to convince Bob in the first attempt, it leaves no room for further persuasion. Finally, the example highlights that the sender's choice of information structure can shape the dynamics of the persuasion process. Only when Alice chooses B in her first persuasion attempt will the persuasion process be sequential. To formalize these insights, we develop the general model in the following section.

3 Model Setup

Our model examines a scenario where a seller (she) attempts to sell a product with symmetrically uncertain value to a buyer (he) at a fixed price, γ. This scenario can be framed within a Bayesian persuasion context, where the seller is the sender and the buyer is the receiver. The model involves two possible states, $\Omega = \{h, l\}$. The product value is 1 when $\omega = h$ and 0 otherwise. The sender's payoff, $u = \alpha$, depends solely on the receiver's action, $\alpha \in \mathcal{A} = \{0, 1\}$, independent of the actual state. If the receiver's belief does not support the sender's preferred action, $\alpha = 1$, the sender can influence the receiver's belief with information structures $\pi : \Omega \to \Delta\Theta$, which map the states to a full support distribution over possible signals. Upon receiving a signal, the receiver updates his belief according to Bayes' rule, and makes a final decision on the action when no further information will be provided. The receiver's payoff, $v = \alpha[\mathbb{1}(\omega = h) - \gamma]$, where $\gamma \in (0, 1)$, is determined by both his action and the actual state.

The sender and receiver share an initial prior belief, $q_1 = \text{Prob}(\omega = h) < \gamma$, about the state. Consequently, if the sender opts not to conduct any experiment, the receiver will choose $\alpha = 0$ based on this prior belief.

Based on this framework, we modify the information transmission phase in the persuasion game to be endogenously sequential, allowing for up to $T \in \mathbb{Z} \bigcap [2, \infty)$ periods. Initially, we assume $T = 2$ to investigate a two-stage persuasion game, then relax this restriction in subsequent sections to examine general intertemporal problems. In each period where $t \leq T$, the sender can propose an experiment defined by the information structure, π_t, and commit to sending the signal generated by π_t truthfully, provided that the receiver agrees to reveal the signal. The receiver benefits from additional information from the experiments, which allows him to make a more informed decision. However, revealing the signal incurs costs. We define the receiver's motivation $\mathcal{M}_t$ in period t as a time-invariant function of the prior belief q_t and the information structure π_t, as follows:

$$\mathcal{M}_t := \mathcal{M}(q_t, \pi_t) = \sum_{\theta \in \Theta} \max\{0,\ \pi_t(\theta|h) q_t - \gamma\} - c(q_t, \pi_t), \tag{1}$$

where π_t is represented by $\pi_t(\theta|\omega)$: the probability of the state $\omega \in \Omega = \{h, l\}$ sending a signal $\theta \in \Theta$ in the period t. The receiver agrees to reveal the signal in the period t only when the information structure proposed by the sender results in $\mathcal{M}(q_t, \pi_t) \geq 0$.

In this motivation function, the first term is the benefit that the receiver expects to obtain, calculated based on the probability of each signal, and the expected payoff conditional on the realization of these signals. The second term in (1) represents the cost that the receiver has to pay (to Nature) to reveal the signal. In this section, we only assume that the cost function is non-negative and continuous in q_t, deferring further specification until the next section, where we discuss the model simplification. This allows us to discuss the properties of the cost function in relation to the receiver's posterior belief distribution, which ensures the model's tractability.

With these settings, the game proceeds as follows: In period 0, Nature selects the state, which is unknown to both sender and receiver. They then form a common prior belief, q_1.

Starting in Period 1, the game enters the information transmission phase. In each period, the sender proposes an information structure, which the receiver observes. Upon observing this structure, the receiver decides whether to incur the cost to reveal the signal. If he decides not to reveal it, the period ends without any information transmission, leaving prior beliefs unchanged for all players. If he chooses to reveal the signal, the signal θ_t is publicly drawn according to the sender's proposed information structure, and both players update their beliefs based on this realization. Should the information transmission phase continue, this updated belief becomes the prior belief for the next period.

The information transmission phase terminates either when the attempt limit T is reached or if the sender chooses to stop proposing information structures. The sender does so if there is no information structure that meets the condition $\mathcal{M}_t \geq 0$ given the prior belief in period t, or if the receiver's prior belief supports his choice of action 1, rendering further persuasion unnecessary. Once the information transmission phase ends, the game enters a decision phase, where the receiver must make a final decision, α. Then the game terminates and payoffs are realized. The timing of the persuasion game is shown in the figure below.

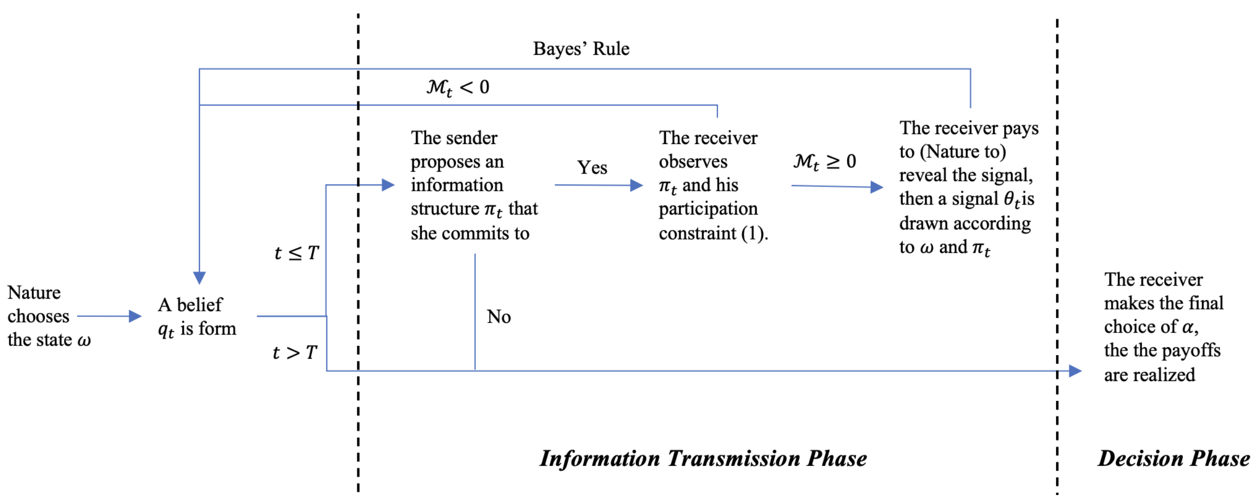

Figure 1: Timeline of an Endogenous Sequential Bayesian Persuasion Game

The timeline above indicates that the sender cannot predict the signal realization and thus does not know ex-ante when she should terminate the information transmission phase.

However, previous research and the following sections of this paper demonstrate that in a binary-state, binary-action setting, the sender's optimal information structure in each period can only generate two signals, either recommending the receiver to act or not. Therefore, the sender can design an intertemporal persuasion path ex-ante, with the decision to terminate the information transmission phase contingent on the signal realization. This persuasion path is characterized by a sequence $(q_1, ...q_\tau)$, where each q_t represents the prior belief assuming the information transmission phase continues with previous signals keeping recommending "not act." For each node along this path, the sender chooses the optimal information structures, which altogether minimizes the overall probability of no signal recommending "act." The sender's objective throughout the information transmission phase is therefore structured as follows:

$$\min_{\tau \in \mathbb{Z} \bigcap [1,T]} \min_{\pi \in \Pi^\tau} \prod_{t=1}^{\tau} \sum_{\theta \in \Theta} \left[\pi_t(\theta|h) q_t + \pi_t(\theta|l)(1-q_t) \right] \mathbb{1} \left(\frac{\pi_t(\theta|h) q_t}{\pi_t(\theta|h) q_t + \pi_t(\theta|l)(1-q_t)} < \gamma \right). \quad (2)$$

In this optimization problem, the sender chooses a series of information structures from the feasible set $\Pi^\tau = \Pi_1 \times ... \times \Pi_\tau$, which satisfies both Bayes plausibility and the frictional constraint $\mathcal{M}_t \geq 0$ due to the receiver's inattention. Additionally, the sender optimizes the length of the persuasion path, τ. While previous studies generally treated τ as a predetermined parameter and focused on how the sender chooses $\{\pi_t\}$, $t = 1, ..., \tau$, here the selection of τ, shaped by the frictional constraint, is an endogenous decision. It distinguishes this study as an endogenous sequential Bayesian persuasion (ESBP) game, setting it apart from previous conventional models.

We use Subgame Perfect Equilibrium (SPE) as the equilibrium concept for this sequential game. With a finite T, even as it approaches ∞, there is always a final period in the information transmission phase, in which the sender will make a take-it-or-leave-it offer for the information structure, and the receiver will choose to accept. When the frictional constraints in each period t are binding, the receiver is indifferent between different information structures proposed across stages. Accepting any proposal that satisfies the constraint becomes his best response. Thus, the equilibrium can be determined using backward induction.

4 Simplification

To make the endogenous sequential Bayesian persuasion game tractable, we start by simplifying the sender's problem and then establishing an analytical framework on this basis. As suggested by Kamenica and Gentzkow (2011), the sender's Bayesian persuasion problem can be reduced to choosing the distribution of the receiver's possible posterior beliefs. We also follow this approach to simplify our model. In a conventional static model with binary states and binary actions, this simplification suggests that the sender will optimally choose an experiment that recommends the receiver to either "act" or "not act." Let $g, b \in \Theta$ denote the good and bad signals recommending "act" and "not act," respectively. Correspondingly, p_t^g and p_t^b denote the posterior beliefs regarding $\omega = h$ in period t. Then a non-dominated, Bayes-plausible strategy, $p_t = (p_t^b, p_t^g)$, requires that $p_t^g \geq \gamma$ and $p_t^b \leq q_t$.

The sender faces a static persuasion problem in the final attempt of a sequential persuasion effort. In this final period t, she chooses the support of the posterior belief distribution, $p_t = (p_t^b, p_t^g) \in [0, q_t] \times [\gamma, 1]$, to maximize the persuasion value conditional on the prior belief q_t. Given p_t and q_t, the posterior belief distribution, $\mu_t = (\mu_t^b, \mu_t^g) \in \Delta\Delta\Omega$, is uniquely determined, with μ_t^θ representing the probability of reaching the posterior p_t^θ. Thus, the sender's information structure, π_t, can be fully characterized by p_t. Accordingly, her optimization problem in this final stage t becomes:

$$\begin{aligned} &\min_{p_t \in [0,q_t) \times [\gamma,1]} \frac{p_t^g - q_t}{p_t^g - p_t^b} \\ &s.t.\ \mathcal{M}(p_t, q_t) \geq 0. \end{aligned} \tag{3}$$

Lemma 1. *For any given $q_t \in (0, \gamma)$, if $\mathcal{M}(p_t, q_t)$ is differentiable in p_t for all $p_t \in [0, q_t] \times [\gamma, 1]$, then there exists an optimal strategy $p_t^* \in [0, q_t] \times [p, 1]$ for the problem (3).*

The continuity of $\mathcal{M}(p_t, q_t)$ in p_t ensures that the set $\{p_t | \mathcal{M}(p_t, q_t) \geq 0\}$ is closed with respect to p_t. Combined with the Bayes-plausibility constraint, which guarantees boundedness, this implies that the feasible set of information structures is compact. Given that the objective function is also continuous in p_t, Lemma 1 holds.

Depending on the level of inattention, the receiver may have different levels of motivation to engage in persuasion. This further restricts the feasible set of information structures beyond the conventional constraint, $p_t \in [0, q_t] \times [\gamma, 1]$. When the sender faces both Bayesian plausibility and the receiver's inattention, Lemma 1 indicates the possibility of characterizing an optimal strategy in the sender's final persuasion attempt, given various prior beliefs. Specifically, when $T = 1$, this lemma predicts an optimal strategy for a static persuasion game. The existence of these optimal persuasion strategies helps characterize the sender's optimal strategies preceding her final attempt, using backward induction.

Lemma 2. *In an endogenous sequential Bayesian persuasion game, when $\mathcal{M}(p_t, q_t)$ is quasi-concave in $q_t \in [0, q_1]$, choosing an objective posterior belief p_t^g within $[q_t, \gamma)$ for a persuasion attempt is strictly dominated.*

According to Lemma 2, even prior to the final persuasion attempt, the sender should adopt an experiment that recommends the receiver to either "act" or "not act." Similar to her objective in the final attempt, her goal is to influence the receiver's decision rather than solely adjusting his belief at any node of the designed persuasion path. Therefore, we can use p_t^θ, $\theta = g, b$ to represent the sender's persuasion strategy in any attempts on the persuasion path. The sender's problem can then be reformulated as choosing the distribution of posterior beliefs in each period of the information transmission phase to minimize the probability of "bad" signals appearing in each period. Lemma 2 simplifies the sender's endogenous sequential persuasion problem as follows:

$$\begin{aligned} \min_{\tau \in \mathbb{Z} \bigcap [1,T]} \min_{p_t \in [0,q_t) \times [\gamma,1]} & \prod_{t=1}^{\tau} \frac{p_t^g - q_t}{p_t^g - p_t^b} \\ s.t.\ & \mathcal{M}(p_t, q_t) \geq 0, \end{aligned} \tag{4}$$

where $p_t^g = q_{t+1}$ for all $\tau \in \mathbb{Z} \bigcap [1, T-1]$. This problem is well-defined and predicts the existence of optimality because the strategy space is compact and the objective function is continuous.

Besides the assumptions made in Lemma 1 and Lemma 2 about the properties of $\mathcal{M}(\cdot)$, which set up the sender's simplified problem (4), we need to make a few more basic assumptions about the motivation function to make this research problem interesting. Failure persuasion comes with a cost. Whenever a persuasion attempt fails, the receiver loses patience, and the sender's subsequent persuasion attempts lose effectiveness. We define this nature as **disappointment-penalizing**. Let $p_q^{*\theta}$ denote the optimal static information structure based on the prior belief q. Define $y(p_t^b, q_t)$ as the minimum p_t^g that satisfies $\mathcal{M}_t \geq 0$ conditional on the prior belief q_t and the choice of p_t^b. Additionally, let $\rho(p, q)$ represent $\frac{p^g - q}{p^g - p^b}$, which indicates the chance of a persuasion attempt failing, given that the prior belief is q and the strategy is p. The definitions below specify two different levels of the receiver's motivation in light of the disappointment-penalizing feature.

Definition 1. *The receiver's motivation is **weakly disappointment-penalizing (WDP)** if, given that $q' > q''$, $y(p_t^b, q'') \geq y(p_t^b, q')$ for all $p_t^b \in [0, q'']$. If this motivation also satisfies $\rho(p_{q''}^*, q'') \geq \rho(p_{q'}^*, q'')$, it is **severely disappointment-penalizing (SDP)**.*

When failed persuasion attempts lead to a decay in the prior belief, the sender faces a reduced success rate for each information structure containing a given p_t^b, making the receiver's motivation weakly disappointment-penalizing (WDP). The disappointment penalty can be more severe. When the receiver's motivation is severely disappointment-penalizing (SDP), not only are some effective information structures that were available for previous persuasion attempts no longer feasible, but those information structures that remain feasible become less effective following an unsuccessful persuasion attempt. This definition forms the axiom underlying our subsequent analysis.

Axiom 1. *The receiver's motivation function satisfies Lemmas 1 and 2. Additionally, it is either WDP or SDP.*

Without a disappointment penalty, the receiver's belief can develop counterintuitive "sweet spots," where he becomes more optimistic about future experiments revealing a high state, even if recent evidence suggests otherwise. These "sweet spots" certainly incentivize the sender to prolong the persuasion process. By adding one additional attempt to the persuasion path, she seizes an additional chance to succeed when the good signal appears and gains a better prior belief for the subsequent persuasion if the bad signal is realized. Axiom 1 eliminates these potential belief "sweet spots" and the sender's ad-hoc motivation to prolong the game, which not only aligns the model with the real world more closely but also ensures the research question remains meaningful.

In Kamenica and Gentzkow's (2011) framework, the persuasion value of a static strategy can be visualized as the distance between the value function and the affine hull, as determined by the support of the posterior belief distribution. The simplification in this section extends this idea to an endogenous sequential persuasion context. According to the sender's simplified problem (4), a failure in a persuasion attempt can lead to a subsequent persuasion with an updated prior belief: $q_{t+1} < q_t$. According to the Kamenica-Gentzkow (K-G) approach, the sender's persuasion attempt in period t should be represented as a line segment connecting to the line segment that represents the following persuasion at the position q_{t+1}, and to the value function at the position $y(q_{t+1}, q_t)$. Panels $(a1)$ and $(b1)$ of Figure 2 illustrate the case where $T = 2$. In these diagrams, the line segments representing the sender's first and final persuasion attempts are shown in blue and red, respectively. The distance between the blue line segment and the value functions at position q_1 captures the overall persuasion values of sequential persuasion strategies in a two-stage game.

5 Endogenous Two-Stage Bayesian Persuasion

This section's model assumes $T = 2$. With this simplest setting in our model, the sender is allowed at most one additional persuasion attempt beyond the static persuasion game. She makes a binary choice: whether to remain in a static persuasion game, or extend the game

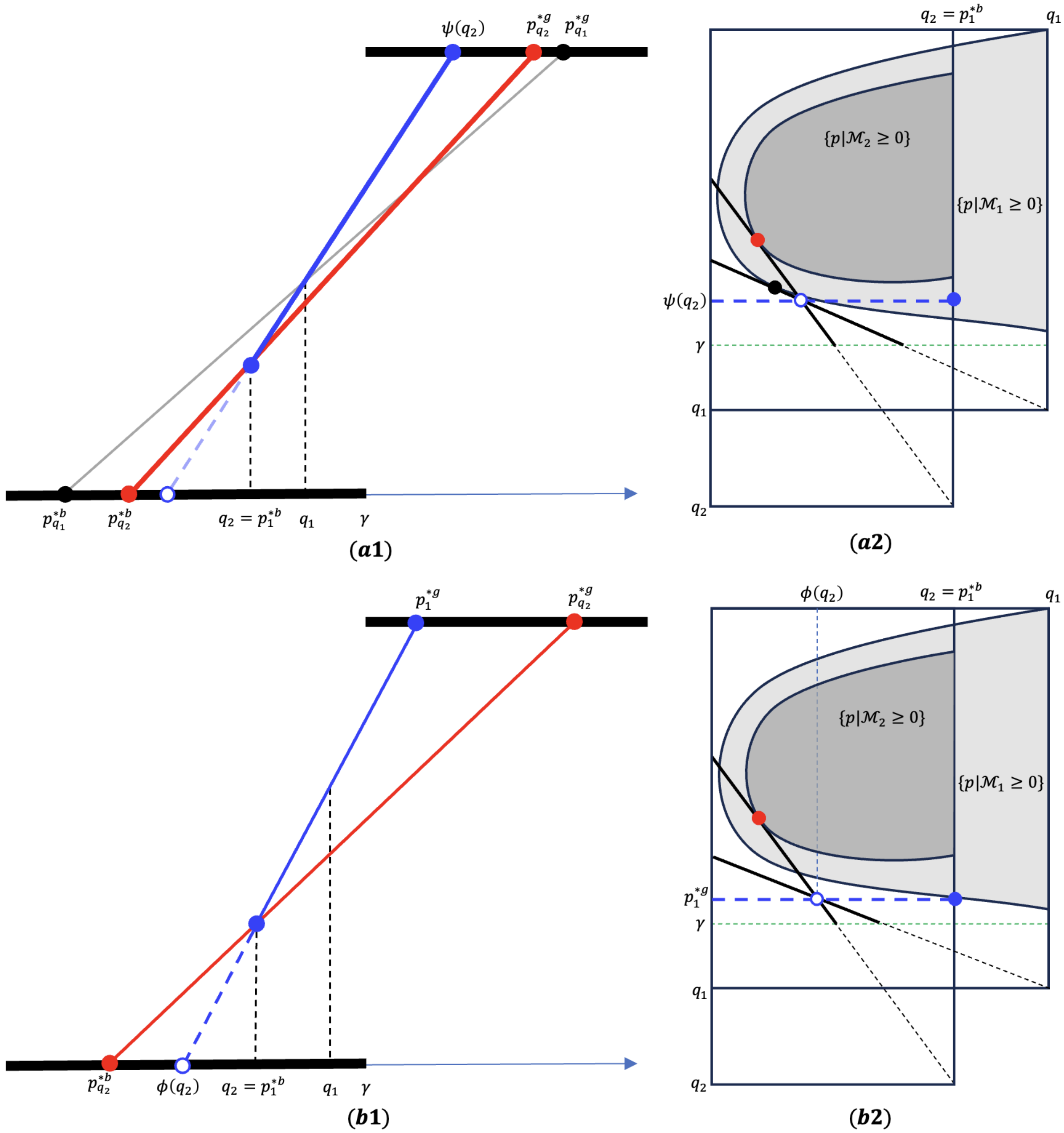

Figure 2: Equivalence Between Strategies with Feasible Sets

to make it sequential. This binary choice allows us to investigate the underlying determinant that makes a sequential game more favorable than a static game for the sender.

5.1 An Additional Persuasion Attempt

With the receiver's varying motivation potentially restricting feasible persuasion strategies, the conventional Kamenica-Gentzkow (K-G) approach is not sufficient to visualize the sender's problem. To make the analysis more intuitive, we develop an alternative framework that characterizes both the set of feasible information structures and their persuasion values. In a two-dimensional coordinate system, where the horizontal and vertical axes represent the receiver's posterior beliefs p^θ given bad ($\theta = b$) and good ($\theta = g$) signals, a feasible set of information structure in period t is $[0, q_t] \times [\gamma, 1] \bigcap \{p_t | \mathcal{M}(p_t, q_t) \geq 0\}$. These sets are shaded regions in panels ($a2$) and ($b2$) of Figure 2. In a static game, the persuasion value of a feasible strategy is represented by $\frac{q_t - p_t^b}{p_t^g - p_t^b}$. On the graph, this value is represented as the slope of the linear indifference curve connecting $p_t = (p_t^b, p_t^g)$ and (q_t, q_t). For a given q_t, a flatter indifference curve indicates a higher persuasion value.

The overall persuasion value of a sequential persuasion strategy is composed of persuasion values from different stages, making it difficult to compare them directly within our framework with indifference curves. To address this challenge, such comparisons must come down to comparing different strategies within a single persuasion attempt. This approach becomes feasible if we can identify either a sequential persuasion strategy equivalent to a static strategy or vice versa.

Lemma 3. *In a two-stage endogenous persuasion game, suppose that the receiver's motivation function satisfies Axiom 1, and the players' common prior belief is q_1 at $t = 1$.*

1) for each $q_2 = p_1^{*b} \in [0, q_1)$, an optimal static persuasion strategy $p_{q_1}^*$ is equivalent to a sequential persuasion strategy where the sender chooses $p_1 = \big(q_2, \psi(q_2)\big)$ and $p_2 = p_{q_2}^*$, provided that $\psi(q_2) \geq \gamma$, where

$$\psi(q_2) = \frac{\rho(p_{q_1}^*, q_1)q_2 - \rho(p_{q_2}^*, q_2)q_1}{\rho(p_{q_1}^*, q_1) - \rho(p_{q_2}^*, q_2)};$$

2) for each $q_2 = p_1^{*b} \in [0, q_1)$, the sender's optimal two-stage persuasion strategy (p_1^*, p_2^*) is *equivalent to a static persuasion strategy,* $p_{q_1} = \big(\phi(q_2), y(q_2, q_1)\big)$*, where*

$$\phi(q_2) = \frac{q_2 - [1 - \rho(p^*_{q_2}, q_2)]y(q_2, q_1)}{\rho(p^*_{q_2}, q_2)}.$$

Proof. Proof of Lemma 3. See Appendix.

Lemma 3 establishes mappings between sets of static and sequential persuasion strategies, indicating their equivalent persuasion values. The first equivalence mapping identifies the first-stage strategies that allow two-stage persuasions with any chosen $q_2 = p_1^{*b}$ to be equivalent to the optimal static persuasion strategy, if such strategies exist. An example of these first-stage strategies appears as a blue line segment in panel $(1a)$ of Figure 2. For this two-stage approach to be equivalent to the static strategy and also lead to the intended q_2, which defines the final-stage persuasion problem, this line segment must attain the persuasion values of the static strategy at position q_1 and the second-stage strategy at position q_2. Accordingly, in panel (a2), $\psi(q_2)$ is determined by the intersection of the indifference curves representing $1 - \rho(p^*_{q_1}, q_1)$ and $1 - \rho(p^*_{q_2}, q_2)$. This equivalence transforms the optimal static persuasion strategy into a series of benchmark sequential persuasion strategies indexed by q_2. To evaluate whether a sequential strategy outperforms the optimal static persuasion strategy, comparing feasible first-stage strategies with these benchmarks is sufficient.

From a different perspective, the second half of Lemma 3 translates optimal sequential strategies, conditional on different choices of $q_2 = p_1^{*b}$, into equivalent static strategies. To find such static strategies, the blue line segment representing the first-stage strategy is extended to make its both ends connect to the value function curve in panel $(b1)$. In panel $(b2)$, this equivalent static strategy is identified by projecting the point representing the first-stage strategy leftward onto the indifference curve for $1 - \rho(p^*_{q_2}, q_2)$. This equivalence allows optimal sequential strategies conditional on different q_2 to be directly compared as static strategies.

When the receiver is willing to engage in persuasion for a second time as long as his motivation is satisfied, it is the sender's decision that determines whether the persuasion is static or sequential. To make the optimal decision in this endogenous sequential persuasion problem, the sender must evaluate whether sequential or static persuasion optimizes her objectives. The analytical framework and the equivalence mappings established in Lemma 3 simplify the comparison between sequential and optimal static strategies, providing a straightforward method to determine which is preferable.

Proposition 1. *Define a benchmark function as follows:*

$$r(q_2) = \begin{cases} \frac{q_1 - \rho(p^*_{q_1}, q_1) q_2}{1 - \rho(p^*_{q_1}, q_1)} & \text{if } \{p_2 | \mathcal{M}\big(p_2, q_2\big) \geq 0\} = \emptyset \\ \psi(q_2) & \text{if otherwise} \end{cases}.$$

Sequential persuasion outperforms static persuasion for the sender if and only if there exists a $q_2 \in [0, q_1)$ *such that* $y(q_2, q_1) < r(q_2)$.

Proof. Proof of Proposition 1. See Appendix.

In the subgame equilibrium, the sender will always choose the optimal static persuasion strategy, $p^*_{q_2}$, for her final attempt, provided that the persuasion game reaches the second stage with prior belief q_2. Following backward induction, her problem in the first stage of persuasion design reduces to choosing the optimal subsequent prior belief q_2 for a potential second attempt. This choice determines both the persuasion value of her possible second attempt, $1 - \rho(p^*_{q_2}, q_2)$, and the probability of success at the first stage, $\frac{q_1 - q_2}{y(q_2, q_1) - q_2}$. To maintain status quo when switching from the optimized static persuasion approach to a sequential one, the sender should ensure that the success probability for her first attempt remains sufficiently high for each possible q_2. $r(q_2)$, which is defined in Proposition 1, indicates these required probabilities. If there exist feasible strategies where p^g_1, lower bounded by $y(q_2, q_1)$, is smaller than $r(q_2)$, the likelihood of first-attempt success increases without reducing the persuasion value of a possible second attempt. As a result, the sequential strategy outperforms the optimal static strategy.

Proposition 1 reduces the comparison between optimal static and sequential persuasion strategies to comparisons between feasible first-stage strategies and the benchmark. Thus, the general question of whether a static or sequential approach is preferable can largely be inferred from the properties of the functions $r(q_2)$ and $y(q_2, q_1)$ under different conditions. When $p^*_{q_1} = (0, \gamma)$, the shape of $r(q_2)$ is pinned down, which establishes Proposition 2 as the baseline case where the receiver is fully attentive to the persuasion.

Proposition 2. *If the strategy $p_{q_1} = (0, \gamma)$ is always feasible within the static persuasion approach, then the sequential persuasion approach never outperforms it. If $\mathcal{M}(p_t, q_t) \geq 0$ always holds, indicating that the receiver is willing to pay attention to any experiment with a Bayes-plausible information structure, then the persuasion value of an optimal static strategy is always equivalent to that of an optimal sequential strategy.*

Proof. Proof of Proposition 2. See Appendix.

Proposition 2 presents two arguments. They are two sides of a coin but have different important implications. First, if Bayes-plausibility is the only constraint on feasible information structures, the sender will choose $p_2 = (0, \gamma)$ whenever the persuasion game reaches the second stage, regardless of the second-stage prior belief q_2. Fixing this optimal subgame strategy, the sender can increase q_2 to improve her expected payoff at $t = 2$; however, this change reduces her payoff at $t = 1$. Proposition 1 shows that the benefits and costs associated with this belief shift offset each other perfectly, as reflected by $\phi(q_2) \equiv 0$, according to the statement 2) of Lemma 3. When there is no predetermined dynamic in the persuasion game, and neither information cost nor discounting is considered, the sender remains indifferent between a static and sequential approach. Therefore, it is justified to conduct research on persuasion within a static framework, without sacrificing generality, as in previous conventional studies.

Second, even if the receiver does not pay attention to all Bayes-plausible persuasion strategies, as long as $(0, \gamma)$ remains a feasible information structure, it becomes the sender's optimal strategy in a static persuasion game. In this case, $r\big(q_2\big) \equiv \gamma$, it is always true that $r\big(q_2\big) = \gamma \leq y(q_2, q_1)$. Therefore, to motivate a sender to choose a sequential persuasion approach over a static one, it is necessary not only that the receiver is inattentive but also that this inattention renders the optimal frictionless static persuasion strategy infeasible. This incentive incompatibility between the sender and the receiver is necessary to trigger the intertemporal trade-off, which causes the sender to consider persuading sequentially. A more important implication of this proposition is that a frictional constraint, which restricts information structures beyond the Bayes plausibility, does not inherently make the sender favor a sequential persuasion approach. This implication emphasizes the importance of analyzing the properties of the frictional constraint and how they shape the sender's design of persuasion strategy.

If the information structure $(0, \gamma)$ is infeasible due to the frictional constraint, a two-stage persuasion strategy may outperform one that optimizes the static problem. According to Proposition 1, this possibility relies on the comparison between $r(q_2)$ and $y(q_2, q_1)$, which are both determined by the properties of the motivation function, $\mathcal{M}(p_t, q_t)$. Lemma 3 and Proposition 1 indicate that $\rho(p^*_{q_t}, q_t)$ determines $r(q_2) = \psi(q_2)$, in the cases that $r(q_2) > y(q_2, q_1)$ is possible. For a given objective, the outcome of an optimization problem is a partial measure of the constraint. Accordingly, how $\mathcal{M}(p_t, q_t)$ varies in the receiver's prior belief q_t is partially measured by $\rho(p^*_{q_t}, q_t)$ and is thus relevant to $r(q_2)$. $y(q_2, q_1)$, on the other hand, measures the lower boundary of the feasible information structures set for a given prior belief q_1. It reflects how $\mathcal{M}(p_t, q_t)$ varies in the information structure p_t.

How the motivation function $\mathcal{M}(p_t, q_t)$ varies in p_t and q_t are independent for certain values of q_2, especially when it is distinct from q_1. As a result, $r(q_2)$ and $y(q_2, q_1)$, both functions of q_2, are generally independent. This observation directly leads to the corollary that follows.

Corollary 1. *Among all motivation functions satisfying Axiom 1 that yield $y(q_2, q_1)$ such that $\{q_2 \in (0, q_1) | y(q_2, q_1) > \gamma\} \neq \emptyset$, there exist some that make a sequential persuasion strategy outperform the optimal static persuasion strategy.*

Corollary 1 joins Propositions 1 and 2 to constitute a general answer to the question of whether and under what conditions an endogenously sequential approach can outperform the static one in persuasion. When frictional constraints arise due to the receiver's inattention, a static framework is insufficient to fully capture a sender's persuasion strategy. Introducing a framework that allows the sender to optimally choose sequential persuasion is necessary to understand insights in the persuasion behavior. Moreover, this corollary suggests that, when certain conditions are satisfied, we can always identify motivation functions that support the sender's choice to extend the persuasion process. These conditions are general and do not preclude the natural situation where certain punishment is imposed based on a history of failed persuasions. When Axiom 1 is considered and that the motivation is either SDP or WDP, the sender will still have incentive to persuade sequentially when these conditions are satisfied.

5.2 Optimal Persuasion Strategy with the SDP Motivation

It is important to investigate the sender's optimal persuasion strategy under the condition that the receiver's motivation is disappointment-penalizing. Since it is natural for the receiver to become less patient after being disappointed by previous unsuccessful persuasion attempts, the sender's decision of extending the persuasion game is justified only if it withstands the receiver's disappointment-penalizing motivation. Additionally, the disappointment penalty imparts specific characteristics to the motivation function, enabling a more intuitive analysis of the conditions that motivate a sequential approach over a static one. This analysis reveals the underlying reasons that sequential persuasion can outperform static persuasion.

Proposition 3. *If the motivation function is SDP and* $(0, \gamma)$ *is infeasible for any* $q_t \in [0, q_1]$, *then a sequential persuasion strategy outperforms the optimal static strategy only when* $p_1^{*b} = q_2 > p_{q_1}^{*b}$ *is chosen and there exists a* $q_2 \in (p_{q_1}^{*b}, q_1)$ *such that* $y(q_2, q_1) < p_{q_1}^{*g}$.

Proof. Proof of Proposition 3. See Appendix.

The SDP motivation function implies that applying the optimal static persuasion strategy in the first attempt leaves no opportunity for further persuasion if the initial attempt fails. This approach is aggressive: upon receiving a "bad" signal, the receiver becomes almost certain of the low state. At this point, only a "good" signal from a highly informative experiment, which incurs significant information costs, could potentially change his mind. As a result, he will not pay attention to a second persuasion attempt. Alternatively, the sender may choose a p_1^b above $p_{q_1}^{*b}$. This less aggressive strategy renders a "bad" signal "less bad," preserving the possibility of a second persuasion attempt if the first one fails.

Since $q_2 < p_{q_1}^{*b}$ would close off future persuasion opportunities, the sender must set q_2 above $p_{q_1}^{*b}$ to secure the opportunity for a potential second attempt. With an SDP motivation function, $\psi(q_2)$ is smaller than $p_{q_1}^{*g}$ for all $q_2 > p_{q_1}^{*b}$. It implies that if $y(q_2, q_1)$ stays at $p_{q_1}^{*g}$ when the sender increases q_2 above $p_{q_1}^{*b}$ to secure an additional persuasion opportunity, this additional chance cannot fully offset the loss in first-stage success probability. Therefore, for a sequential persuasion strategy to outperform the static approach, it is necessary that a $p_1^{*b} = q_2$ larger than $p_{q_1}^{*b}$ is paired with a $p_1^{*g} = y(q_2, q_1)$ smaller than $p_{q_1}^{*g}$. This ensures that a larger q_2 does not cost too much success rate loss in the sender's first persuasion attempt.

According to Proposition 3, whether sequential persuasion outperforms static persuasion is determined by the trade-off between the additional persuasion value in the second attempt and the persuasion value that can be lost in order to create a second chance to persuade. A higher penalty for previous failure requires a stronger negative correlation between $y(p_t^b, q_t)$ and p_t^b to ensure that the persuasion value lost is outweighed by the additional value created. For example, if $\frac{\partial \rho(p,q)}{\partial q} > 0$ is marginally increasing—representing an even more severe penalty for persuasion failure than that suggested by an SDP motivation function—then $r(q_2)$ decreases across all $q_2 \in [p_{q_1}^{*g}, q_1]$. According to Proposition 1, only a significant negative correlation between $y(p_t^b, q_t)$ and p_t^b would be enough to ensure the existence of a $q_2 = p_1^b$ such that $y(q_2, q_1) < r(q_2)$.

5.3 Optimal Persuasion Strategy with the WDP Motivation

With SDP motivation, if the sender fails to modify her conventional static strategy to allow for a potential second persuasion attempt, she will completely miss this second opportunity. This stark difference in outcomes—depending on whether or not the strategy is modified—may encourage the sender to pay closer attention to avoid such an oversight. However, this distinction becomes more subtle and easier to miss when the motivation function is WDP rather than SDP, reflecting a less severe penalty for failed attempts. Although $y(p_t^b, q_t)$ increases as q_t decreases with SDP motivation, if the increase is not substantial for $p_t^b \leq p_{q_1}^{*b}$, setting $q_2 = p_{q_1}^*$ can still produce a chance for subsequent persuasion. In this case, some sequential persuasion strategies definitely outperform the optimal static strategy, as the sender gains additional persuasion value without losing any chance in her first persuasion attempt.

However, it remains possible that the optimal static persuasion strategy is not the best choice for the first attempt in a two-stage sequential persuasion design. If the sender fails to recognize that sequential strategies can outperform the optimal static strategy and design the strategy accordingly, she may adhere to the optimal static strategy in her initial attempt. While she may still have a second chance to persuade, her suboptimal initial strategy would cause her to miss the opportunity to implement the most effective persuasion strategy.

Proposition 4. *Given that the motivation function is WDP and supports at least one sequential persuasion strategy that outperforms the optimal static persuasion strategy, the optimal sequential strategy indicates that the sender should optimally choose $p_1^{*b} > p_{q_1}^{*b}$ when the disappointment penalty is sufficiently large.*

Proof. Proof of Proposition 4. See Appendix.

Since $\mathcal{M}(p_t, q_t)$ is differentiable in p_t, the sender's first order condition (F.O.C.) of optimizing a sequential persuasion problem is given as follows:

$$[y(p_2^b, q_2) - q_2]\Lambda + \frac{(q_2 - p_2^b)y'_{q_2}(p_2^b, q_2) + [p_2^b - y(p_2^b, q_2)]}{y(p_2^b, q_2) - p_2^b} \left(\frac{y(q_2, q_1) - q_1}{y(q_2, q_1) - q_2} \right) = 0, \tag{5}$$

where $y'_{q_2}(p_2^b, q_2) < 0$ represents the derivative of $y(p_2^b, q_2)$ with respect to q_2, which measures the disappointment penalty, and Λ is the first order derivative of the objective function $\rho(p_{q_1}, q_1)$ with respect to $p_{q_1}^b$.

As $p_{q_1}^{*b}$ minimizes the objective $\rho(p_{q_1}, q_1)$, Λ is less than or equal to 0 when p^b approaches $p_{q_1}^{*b}$

from the left, and greater than or equal to 0 from the right. However, with a WDP motivation and $y'_{q_2}(p_2^b, q_2) < 0$, a larger prior belief q_2 in the second persuasion attempt not only improves the success rate of each information structure but also lowers the disappointment penalty. In the absence of a trade-off, this shift in belief, q_2, always yields a higher persuasion value for the second attempt. If the punishment, as measured by $|y'_{q_2}(p_2^b, q_2)|$, is sufficiently large, then the left-hand side of (5) becomes negative at $p_{q_1}^{*b}$, leading to $p_1^{*b} > p_{q_1}^{*b}$.

Given the receiver's inattention, his disappointment-penalizing motivation does not prevent the sender from planning a sequential persuasion strategy. However, if she anticipates a potential second attempt, she is reluctant to shift the receiver's belief drastically in the first attempt. This conservative approach allows her to secure a second persuasion opportunity under SDP motivation and achieves a higher subsequent persuasion value even if only WDP motivation is present.

In brief social interactions, making an immediate strong impression can be challenging. While advice often emphasizes making a good first impression, it can overlook both the difficulty of doing so and the potential consequences of a failed attempt. The listeners may not have enough time to fact-check an assertive introduction, and if they remain unconvinced, the introduction may come across as boastful. Therefore, unless the sender has only one chance to impress, it may be more effective to establish a connection in the social interaction with a more conservative approach. Even if the listeners are not immediately convinced, they are less likely to lose interest in future interactions, providing further opportunities for the sender to make a positive impression over time.

6 ESBP with Binding Attempt Limit

If the sender is allowed additional attempts in the persuasion game beyond two, she may still find it beneficial to design a persuasion path that exhausts the attempt limit. In scenarios where the attempt limit is general but remains binding and finite, the endogenous sequential Bayesian persuasion (ESBP) becomes a general intertemporal problem with T periods for the sender. In each period, q_t serves as the state variable, and $p_t = (q_{t+1}, y(q_{t+1}, q_t))$ becomes the control variable. This framework not only generalizes the findings from the previous section but also enables a detailed analysis of the sender's persuasion path, showing how the information structures and the receiver's potential prior beliefs evolve over time endogenously.

6.1 Additional Attempts in the Interim Subgame Structure

The attempt limit T is binding if the receiver consistently finds that including additional attempts in the sequential persuasion design is beneficial when the number of attempts is below T. To determine whether an additional attempt improves overall persuasiveness, the sender needs to weigh trade-offs in adjacent subgames, as in the two-stage persuasion. However, if the sender can make more than two persuasion attempts, one of these subgames might not be the final one. In other words, the sender's evaluation now includes deciding whether to introduce an additional interim attempt, where failure would lead to yet another interim persuasion stage, rather than the final attempt.

Proposition 5. *Consider a sequential strategy involving adjacent subgames Q_t and $Q_{t+1} \subset Q_t$, with prior beliefs q_t and $q_{t+1} < q_t$, respectively. For each belief $q' \in (q_{t+1}, q_t)$, adding an additional attempt in the persuasion path is beneficial for the sender if and only if the following holds: when this additional attempt causes $y(q_{t+1}, q_t)$ to increase to $y(q_{t+1}, q')$ by $\delta_{t+1} > 0$, it results in $y(q', q_t)$ being less than $y(q_{t+1}, q_t)$ by more than*

$$\delta_t = \frac{(q' - q_{t+1})\rho(q_{t+1}, q_t)\delta_{t+1}}{q_t - q' + (1 - \rho(q_{t+1}, q_t))\,\delta_{t+1}}.$$

Proof. Proof of Proposition 5. See Appendix.

As discussed in the previous section, the optimal strategy for each persuasion attempt yields two direct recommendations for the receiver, "act" and "not act." Only under the latter recommendation does the sender have an incentive to continue persuading. Hence, given the receiver's motivation, the sender's sequential persuasion design can be characterized as a series of prior beliefs following each failed attempt, $Q_\tau = (q_1, ...q_\tau)$, $\tau \leq T$. This series essentially represents a Markov Chain, with the probabilities of recommendations in each stage and their associated expected payoffs determined solely by prior belief from the previous stage.

Given the properties of the persuasion process, adding a persuasion attempt to an interim two-stage subgame structure affects only that structure, impacting overall persuasiveness solely through its influence on this component. Specifically, the additional attempt directly influences the prior belief in the persuasion stage that immediately follows it. Here, δ_{t+1}, defined as the change in $y(q_{t+1}, \cdot)$ due to a shift in prior belief from q_t to q', indicates how a less favorable prior belief ($q' < q_t$) changes the maximum probability of success in the following attempt to persuade. With a less favorable prior belief, a more informative signal is required to push the receiver's belief across the cutoff γ and prompt the receiver to change his decision to "act." A positive δ_{t+1} aligns with Axiom 1, which characterizes the disappointment-penalizing nature of the receiver's motivation to engage with persuasion. This implies that the deteriorated prior belief makes the original persuasion design prohibitively costly for the receiver to engage with. Consequently, the sender must make the "good" signal much better to regain the receiver's attention. This adjustment in the strategy design lowers the success rate of the attempt immediately following the inserted stage.

When a failed persuasion attempt disappoints the receiver, the persuasion path will only include a new attempt if it produces additional chances of success. This requires a smaller p^g in the persuasion attempt preceding (q_{t+2}, q_{t+1}). In Proposition 5, a positive δ_{t+1} implies a positive δ_t, which requires that $y(p_t^b, q_t)$ decreases with p_t^b. Although only WDP is assumed, the requirement for adding an additional attempt in the interim subgame structure resembles that in Proposition 3, which assumes SDP. When the next persuasion attempt is the final one, the sender can adjust p_t^b to compensate for the receiver's disappointment when only WDP is assumed. However, if this additional persuasion attempt leads to another interim attempt, p_t^b is impacted by the following persuasion path, thereby losing certain degrees of freedom. This limitation necessitates a greater compensation from the inserted persuasion attempt.[4] Generally, if the receiver's motivation is disappointment-penalizing, which indicates

[4]In Proposition 5, p_t^b is fixed at q_{t+1}. But as explained below, Proposition 5 is only a sufficient criterion for adding persuasion attempts, which implies that the sender may adjust the persuasion path following $t+1$

that $y(p_t^b, q_t)$ increases in q_t, the motivation's property that leads $y(p_t^b, q_t)$ to decrease in p_t^b is favorable for encouraging sequential persuasion, particularly when T is larger. Certain information costs, formulated as functions of mutual information, exhibit this property.

The overall impact of adding a persuasion attempt within an interim subgame structure can be represented within the extended framework from Figure 2. Panel (b) of Figure 3 depicts a scenario where at most three persuasion attempts are permitted ($T = 3$). Given the sequential strategy in this graph, the process of determining the equivalent static persuasion strategy begins with the top endpoint, which represents the optimal strategy in the final attempt. From here, the process moves downward along iso-value curves, transitioning to a different curve at each connecting point until reaching the bottom endpoint. This process is equivalent to pivoting the line segments from left to right around these connecting points (from q_2 to q') in panel (a), effectively raising the overall position of the line segment configuration (indicated by the elevation of the red line segment).

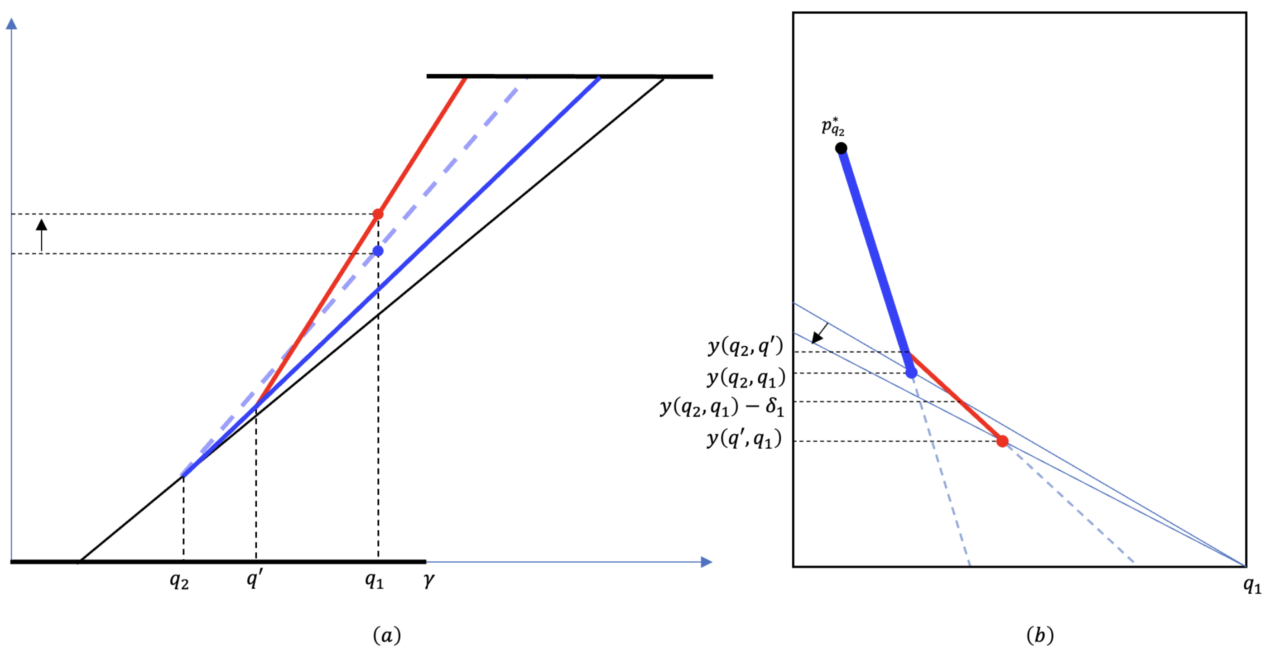

Figure 3: Value of an Additional Persuasion Attempt $(q', y(q', q_1))$

to lower the disappointment penalty and accommodate for the additional persuasion attempt. However, this adjustment still involves more than one stage of persuasion design. Therefore, the degree of freedom in the adjustment is still restrictive. See Appendix for a more detailed discussion

This extended framework provides a rudimentary yet intuitive approach to designing a sequential persuasion strategy. The sender begins by identifying a starting point in the upper left corner of panel (b), which is then connected to a sequence of line segments representing the indifference curves of prior beliefs along the persuasion path. These segments, with decreasing slopes in order, link together to form a "wire" that transmits the starting point to the lower right, ultimately reaching the receiver's indifference curve associated with q_1 at the lowest position.

Based on this intuition, Figure 3 implies how adding a persuasion attempt, even within an interim subgame structure, can improve the overall persuasiveness. In panel (b), an additional red line segment causes the connecting point to retreat to a higher position, $y(q_2, q')$. However, as long as the red segment reaches below the benchmark level, $y(q_2, q_1) - \sigma_1$, the entire "wire" with this additional component reaches a lower indifference curve, indicating improved overall effectiveness. Panel (a) likewise reflects this improvement.

In Proposition 5, the value of $\rho(q_{t+1}, q_t)$ is unspecified, and the additional stage to be inserted can represent the combination of multiple attempts. Thus, this proposition can serve as a very general rule of thumb for the sender when designing a persuasion path, helping her decide whether to expand the persuasion path to improve persuasiveness, no matter which path she is currently considering. For example, if q_t and q_{t+1} belong to the optimal persuasion path with τ attempts, this proposition becomes a sufficient condition under which the sender can improve her ex-ante persuasiveness by extending the design of the persuasion path beyond τ attempts.

According to Proposition 5, whether or not to include an additional attempt within the interim subgame structure depends on the comparison between the benchmark δ_t and the actual variation of $y(\cdot, q_t)$. Given a persuasion path, the benchmark is determined by a series of $\{y(p^b, q)\}$, where $p^b = q_{t+1}, ...q_\tau, p^{*b}_{q_\tau}$ and $q = q_t, q_{t+1}, ..., q_\tau$. Since in $y(p^b, q)$, $p^b < q'$ when $q = q_t$, the benchmark δ_t and the variation of $y(\cdot, q_t)$ are independent. This directly produces Corollary 2.

Corollary 2. *For a motivation function $\mathcal{M}(p_t, q_t)$ where $Q_\tau = (q_1, ...q_\tau)$ is an optimal persua-*

sion strategy under the condition that the total attempts is $\tau < \infty$, there exists an alternative motivation function $\tilde{\mathcal{M}}(p_t, q_t)$ such that an extended persuasion strategy $\tilde{Q}_{\tau+1} = (q_1, ., q'..q_\tau)$ generates a higher overall persuasion value than Q_τ, with $\mathcal{M} = \tilde{\mathcal{M}}$ at $q_t = q_1, ...q_\tau$.

Corollary 2 suggests that any finite-length persuasion path can be supported by certain motivation functions. It also highlights the sensitivity of the path length to the receiver's inattention-based motivation. A small modification of the receiver's information cost function within the range $[q_{t+1}, q_t]$ can substantially change the optimal length of the persuasion path. Therefore, just because the sender prefers sequential persuasion over static persuasion does not inherently imply that she prefers more attempts in a sequential persuasion strategy in general. On the other hand, a two-stage sequential persuasion design may be less effective than the optimal static persuasion attempt, but adding more stages could ultimately lead to a sequential strategy that outperforms the static approach in overall persuasiveness. Therefore, while a two-stage persuasion game provides intuitive insight into the endogenous mechanism of sequential persuasion under frictional constraints, it may not fully capture the sender's optimal path selection, underscoring the need for analysis within the general ESBP framework with $T \leq \infty$ attempts. Moreover, to effectively design an optimal persuasion strategy, the sender should be aware of the limit on persuasion attempts and the nature of receiver inattention, as small changes in conditions can result in significantly different optimal approaches.

6.2 Piecemeal Information Disclosure

Given that the persuasion design includes $T < \infty$ attempts, the sender needs to solve an intertemporal problem to choose optimal information structures along this path. This problem can be formalized using a Bellman function as follows.

$$K(q_t) = \min_{q_{t+1}} \left[ln \left(\frac{y(q_{t+1}, q_t) - q_t}{y(q_{t+1}, q_t) - q_{t+1}} \right) + K(q_{t+1}) \right]. \tag{6}$$

This equation yields the transition (Euler) condition shown below in Lemma 4, which outlines how the optimal information structures evolve in a sequential persuasion strategy from stage 1 to T.

Lemma 4. *Suppose that $Q_T^* = (q_1, ..., q_T)$ represents an optimal persuasion path, $p_t^* = (q_{t+1}, y(q_{t+1}, q_t))$ is the associated optimal information structures at each attempt $t \in [1, T-1]$ with prior beliefs q_t and $q_{t+1} = p_t^{*b}$. The optimal information structures satisfy the relationship indicated by the following Euler condition:*

$$\frac{1}{p_{t+1}^{*g} - p_{t+1}^{*b}} = \frac{1}{p_t^{*g} - p_t^{*b}} + y'_{t,q_{t+1}} \frac{q_t - q_{t+1}}{(p_t^{*g} - q_t)(p_t^{*g} - q_{t+1})} + [y'_{t+1,q_{t+1}} - 1] \frac{q_{t+1} - q_{t+2}}{(p_{t+1}^{*g} - q_{t+1})(p_{t+1}^{*g} - q_{t+2})}, \tag{7}$$

where $y'_{t,q_{t+1}}$ and $y'_{t+1,q_{t+1}}$ denote the partial derivatives of $y(q_{t+1}, q_t)$ and $y(q_{t+2}, q_{t+1})$ with respect to q_{t+1}, respectively.

Proof. Proof of Lemma 4. See Appendix.

Similar to the two-stage persuasion game, the optimal design of a persuasion path in a general ESBP game relies solely on how the motivation function varies with the prior belief q_t and the targeted posterior belief p_t. Based on Lemma 4, specifying these properties allows for deterministic predictions about the evolution of the information structure as the persuasion proceeds along the designed path.

Proposition 6. *Given that $y(q_{t+1}, q_t)$ decreases with both q_t and q_{t+1} for all possible q_t and q_{t+1}, the information structure employed in each period is less informative than information structures employed in following periods, implying that $p_{t+1}^{*g} - p_{t+1}^{*b} > p_t^{*g} - p_t^{*b}$. Specifically, p_t^{*b} decreases and p_t^{*g} increases as t increases from 1 to T.*

Proof. Proof of Proposition 6. See Appendix.

Given that the receiver's motivation is disappointment-penalizing, an information structure with both a larger p^g and p^b than another should be dominated by the latter and, if possible, left out of a persuasion design. If any information structure in the optimal path does not meet the condition $p_{t+1}^{*g} - p_{t+1}^{*b} > p_t^{*g} - p_t^{*b}$ in Proposition 6, then the given attempt limit T

cannot be binding. As a result, the condition $y'_{t,q_{t+1}} < 0$ becomes necessary in this proposition only when the sender must exhaust all allowed persuasion attempts. This condition is also critical for ensuring a positive δ_t in Proposition 5, which motivates the sender to extend the game before reaching the attempt limit. Proposition 6, therefore, emphasizes "piecemeal" information disclosure strategy as a characteristic of the sequential persuasion approach. As long as the sender is motivated to further extend the persuasion game for greater effectiveness, she adopts this "piecemeal" strategy. If she cannot maintain this strategy with more attempts, she ceases to extend the persuasion path, leaving the attempt limit non-binding.

When the sender adopts this "piecemeal" information disclosure strategy, she designs the bad signal to be less bad, ensuring that each potential failure results in only a slight decline in q_t. By doing so, the sender leverages a small p_t^g in each persuasion attempt, while gradually providing more information over time to keep the receiver engaged only when a previous attempt is unsuccessful. Despite the "price" of providing more information for each prior failure, this "piecemeal" strategy allows the sender to maintain the receiver's attention in subsequent attempts. Ultimately, this approach helps the sender achieve a level of persuasiveness that would be unachievable with fewer attempts on the persuasion path.

Che et al. (2023) discussed a scenario where effective signals that lead to belief "jumps" do not arrive immediately (with an arrival rate of $\lambda < 1$ at any time t). According to Proposition 6, such a "jump" may not occur even if the signal arrives each time whenever the experiment is performed. Instead, these less effective signals may be the result of the sender's endogenously optimal persuasion path design, which aims to improve overall persuasion effectiveness by overcoming the receiver's inattention.

This "piecemeal" information disclosure strategy is common in real-world practices, such as advertising campaigns for newly-launched products, government propaganda to shape public opinion, and even job interviews. For example, rarely does a job candidate start the introduction with a detailed and rigorous proof that evidences the contribution of her work, especially in preliminary rounds when time is limited and the interviewer is likely to lack specific background knowledge. Instead, candidates often start their introductions with intuitive examples, which may not fully convey the information necessary to immediately convince the interviewers of the value of their work, but they still manage to keep the interviewers engaged. For this purpose, candidates typically prepare varying lengths of introductions—1-minute, 5-minute, and longer formats—each of which can stand alone, covering similar topics but with different levels of detail. Although this may not be the best way to structure the manuscript for publication, where readers often invest more attention, it is proven highly effective in succeeding in interviews and presentations where the listener's attention is limited.

7 Non-Binding Attempt Limit and Efficiency Boundary

Under certain conditions, structuring the persuasion effort sequentially can mitigate efficiency losses caused by the receiver's inattention. This finding from previous analysis, however, is subject to the binding constraint of a maximum permitted number of attempts. Allowing the sender to persuade as many times as desired can further relax this constraint to improve the effectiveness of persuasion. When T is sufficiently large but finite, the attempt limit becomes non-binding, allowing the sender to continue persuading as long as she can construct information structures that motivate the receiver to engage in persuasion. This relaxation not only shapes the sender's optimal persuasion design but also allows us to characterize the efficiency boundary and understand how much an ESBP game can beat the receiver's inattention.

The binding constraint on the maximum number of attempts in the persuasion game significantly affects the design of the optimal persuasion path. According to Proposition 5, it may be beneficial to include an additional attempt in the persuasion path design. However, if the path has reached the attempt limit, adding an attempt requires replacing an existing one, which could be detrimental to the overall effectiveness. Formally, in the intertemporal persuasion problem, there are many paths that meet the Euler condition (7). However, backward induction optimally narrows the choices to a single terminal subgame, and a path optimally leads to this subgame, leaving out some information structures that could benefit the overall persuasiveness had more persuasion attempts been allowed.

In contrast, if the attempt limit is no longer binding, the trade-off among different designs satisfying the Euler condition vanishes. Any attempt that improves overall effectiveness can be added to the persuasion path without needing to replace an existing one. The following lemma states that if each persuasion attempt has a sufficiently small impact on subsequent stages, the sender's task in designing an optimal persuasion path reduces to determining whether an attempt is qualified for inclusion, rather than selecting among qualified attempts.

Lemma 5. *Suppose there is a persuasion path $Q_\tau = (q_1, ..., q_t, q_{t+1}, q_{t+2}, ..., q_\tau)$ with $\tau < T$. Given a motivation functions with sufficiently small disappointment-penalties for prior beliefs between q_t and q_{t+1}, if there exists an information structure $(q', y(q', q_t))$ that satisfies Proposition 5, the persuasion design with the path $(q_1, ..., q_t, q', q_{t+1}, q_{t+2}, ..., q_\tau)$ outperforms the design with $(q_1, ..., q_t, q', q_{t+2}, ..., q_\tau)$.*

Proof. Proof of Lemma 5. See Appendix.

When the disappointment-penalty, measured by $\left|y'_{t,q_{t+1}}\right|$, is sufficiently small, the receiver's motivation does not substantially vary with his prior belief. This lack of sensitivity to previous persuasion failures means that replacing an information structure in a given attempt has minimal impact on subsequent attempts. Consequently, if the sender finds it beneficial to replace an information structure in the persuasion design, it is generally even more beneficial to append it instead, so that she can take advantage of the chance to succeed in both attempts. With this disappointment-penalizing feature in the receiver's motivation, designing a sequential persuasion strategy can be a process of including attempts that add persuasiveness to the existing path until no additional qualified attempts are available.

For an optimal path with finite attempts, this approach simplifies the sender's task of identifying the optimal persuasion path. Furthermore, this method suggests that an optimal sequential strategy can incorporate as many qualified attempts as exist across other candidate persuasion paths. Within an intertemporal framework, the sender can adjust initial conditions in the Euler equation to ensure that more such attempts are included in the optimal path when no limit is placed on the number of attempts. However, to fully characterize the optimal sequential persuasion design, further specification is required on which types of attempts should be included or excluded from an optimal path.

Definition 2. *With disappointment-penalties in the receiver's motivation and non-binding attempt limits, the information structures* $(q, y(q))$ *and* $(q', y(q'))$ *are* ***mutually dominated*** *in a persuasion path if* $y(q) > y(q')$ *and* $q > q'$*.*

This definition derives from Proposition 5: if mutually dominated information structures appear simultaneously in a sequential persuasion design, the design cannot be optimal, as removing one of these structures will improve the overall persuasiveness. Based on this intuition, we can characterize the sender's optimal sequential persuasion design when the disappointment penalty is sufficiently small. This optimal design is determined solely by the receiver's motivation function.

Proposition 7. *When the attempt limit is not binding, and that the disappointment-penalty is sufficiently small for each possible* q_t*, the optimal sequential persuasion strategy includes all non-mutually dominated information structures* $\left(p_t^b, y(p_t^b, q_t)\right)$*, where* $p_t^b \in [p_\tau^{*b}, p_1^{*b}]$*, assuming there are finite of them. In this design, if* $\left|y'_{t,q_{t+1}}\right| \to 0$*, then* $p_1^{*g} = y(p_1^{*b}, q_1)$ *is the smallest possible* $p^g \geq \gamma$ *that satisfies the participation constraint* $\mathcal{M}_t \geq 0$*, denoted as* p^g_{min}*.* p_τ^{*b} *is the smallest possible* p^b *that satisfies participation constraint* $\mathcal{M}_t \geq 0$*, denoted as* p^b_{min}*.*

Proof. Proof of Proposition 7. See Appendix.

In a binary-state-binary-action Bayesian persuasion problem that is central to our research, the sender should aim to minimize both p^g and p^b to achieve an optimal persuasion outcome. If an information structure within the feasible set includes both p^g_{min} and p^b_{min}, it dominates all alternatives in a static persuasion game. However, the sender's motivation function and its associated constraints may render this dominating strategy infeasible. Although p^g_{min} and p^b_{min} may not coexist on a single feasible information structure, given a sufficiently small disappointment-penalty (i.e., $\left|y'_{t,q_{t+1}}\right| \to 0$), a feasible sequential persuasion design can bridge these endpoints in an endogenous sequential game with a non-binding attempt limit. In this design, the sender can leverage both p^g_{min} and p^b_{min} to maximize the overall persuasiveness.

Therefore, in an ESBP game with a non-binding attempt limit, the key to the sender's strategy design is the selection of two endpoints of the sequential path. According to Proposition 7, once these endpoints are chosen, the sender can search q_t, from q_τ all the way to q_1, including each $(q_t, y(q_t, q_{t-1}))$, with replacement only if the new information structure and any previously included structure are mutually dominated.

Panel (a) in Figure 4 visualizes an example of how a sequential persuasion path, under a non-binding constraint on the maximum attempts, can bridge p^g_{min} and p^b_{min} within the feasible set, given that $\left|y'_{t,q_{t+1}}\right| \to 0$. Starting from the left endpoint, defined by p^b_{min}, non-mutually-dominated information structures extend this "bridge" downward to connect with the lower endpoint defined by p^g_{min}. Graphically, p^g_{min} ensures that the "bridge" extends as low as possible, while p^b_{min} provides the farthest left anchor. Additional non-mutually-dominated interim designs in the sequential strategy reinforce the position of this "bridge," allowing the lower endpoint, which represents the static persuasion strategy equivalent to this sequential strategy, to reach as far to the lower left as possible, thereby enhancing overall persuasiveness.

In Panel (b) of Figure 4, the "bridge" from panel (a) appears as an envelope curve encompassing all line segments that represent the experiments included in the optimal sequential strategy, which altogether form the shaded area on the graph. The elevation of this envelope curve at the position of q_1 reflects the expected effectiveness of the sequential strategy. The graph shows that, while (p^b_{min}, p^g_{min}) is not feasible as any single information structure, the "piecemeal" information disclosure strategy connects the endpoints of the envelope, which approximates the persuasiveness of the static strategy (p^b_{min}, p^g_{min}).

Although the sequential persuasion path can leverage both p^g_{min} and p^b_{min}, they are, after all, not connected within a direct experiment in the static game. Instead, it requires the interim persuasion stages to build up this connection. However, these interim components discount the minimization effectiveness when they connect p^g_{min}, in the first segment, to p^b_{min}, in the

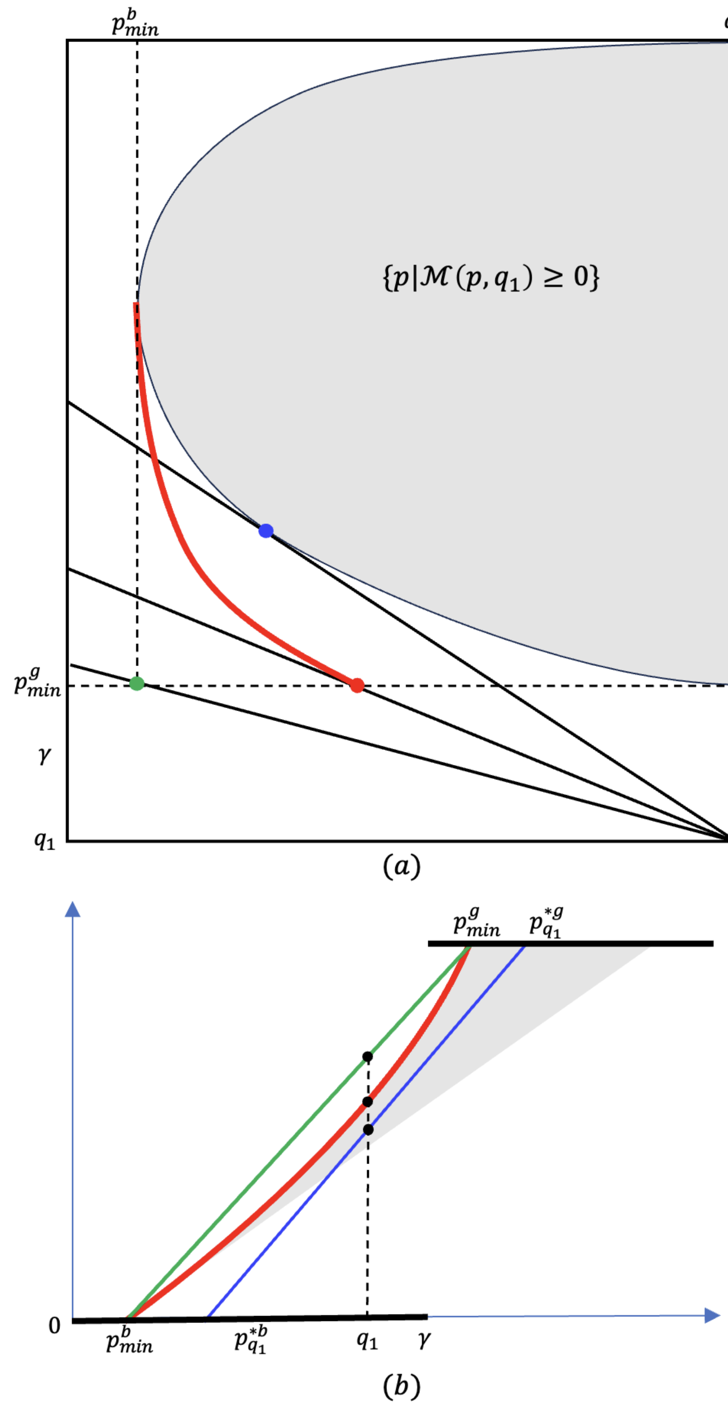

Figure 4: Efficient Boundary of Sequential Persuasion Strategies

last segment, in the chain of $\prod \frac{p^g_t - q}{p^g_t - p^b_t}$. Graphically, the connection between $\left(p^b_{min}, y(p^b_{min}, q_\tau)\right)$ and $\left(q_2, p^b_{min})\right)$ is bowed-out in panel (a) and bowed-in in panel (b) of Figure 4. This curvature implies that the overall effectiveness of the sequential persuasion design, even with $\left|y'_{t,q_{t+1}}\right| \to 0$, is ultimately inferior to a direct experiment in the static game, (p^b_{min}, p^g_{min}).

Corollary 3. *As the disappointment penalty approaches 0, the optimal persuasion path defined in Proposition 7 establishes the efficiency boundary for the endogenous sequential persuasion game. The maximum overall persuasiveness of this persuasion path is bounded above by* $\frac{q_1 - p^b_{min}}{p^g_{min} - p^b_{min}}$.

According to Corollary 3 and its basis, Proposition 7, the efficiency boundary of a sequential persuasion strategy is still determined by the receiver's inattention constraint. With an increased number of permitted attempts and given a "well-behaved" constraint—featuring both a downward-sloping lower boundary and a minimal disappointment penalty—the sender can strategically extend the persuasion process to improve the effectiveness of her attempts. This effort potentially surpasses the optimal outcome in a static persuasion game. However, the frictional constraint still sets the efficiency limit. Specifically, if the frictionless optimal static persuasion strategy, $(0, \gamma)$, is rendered infeasible by the receiver's inattention, it still cannot be achieved through a "piecemeal" information disclosure strategy. Extending the game can push the boundary set by static persuasion under the receiver's inattention, but it cannot fully beat this inattention.

8 Conclusion

This paper discusses an endogenous sequential Bayesian persuasion model where the receiver's rational inattention imposes a frictional constraint on feasible information structures beyond Bayes plausibility. Due to a frictional constraint, sequential persuasion can be more effective than the static approach. This motivates the sender to strategically withhold information to prevent drastic shifts in the receiver's beliefs. Although this "piecemeal information disclosure" strategy may not maximize the probability of success in a single attempt, it preserves the possibility of future attempts if the current one fails, allowing persuasion efforts to accumulate over time, which ultimately outperforms a single, aggressive attempt aimed at immediate success.

We emphasize that in Bayesian persuasion, frictions and sequential processes often coexist, especially when the persuasion game itself is not intrinsically dynamic. This interaction underscores the importance of studying friction and dynamics together in persuasion and even information design research. In this research, we develop and integrate modules into a framework that intuitively analyzes the frictional constraint of information structures and formalizes the sender's problem as intertemporal. This unified approach not only offers insight into the sender's strategic choices but also provides a foundation for future studies to explore additional real-world phenomena.

Current investigations into the receiver's inattention within information design rely on assumptions that are axiomatically consistent with Blackwell's Theorem (1953). In a more general sense, where we do not predetermine this assumption, our analysis reveals a mechanism that links the receiver's rational inattention as an abstract frictional constraint to the sender's strategies. This connection paves ways for future research to estimate the receiver's implicit information cost through observable persuasion strategies and factors like the hazard rate associated with unexpected game termination. Additionally, our findings suggest that the receiver's reluctance to fully engage with an initial persuasion attempt may ultimately cause him to pay more attention to the subsequent persuasion if the sender adopts a sequential strategy. However, this endogenous sequential approach may not be optimal for either the receiver's or overall social welfare. Future research could explore mechanisms that enhance social welfare in persuasion games where the receiver is rationally inattentive.

References

Ashkenazi-Golan, Galit, Penélope Hernández, Zvika Neeman, and Eilon Solan, "Markovian Persuasion with Two States," *arXiv preprint arXiv:2209.06536*, 2022.

Au, Pak Hung, "Dynamic information disclosure," *The RAND Journal of Economics*, 2015, *46* (4), 791–823.

Bizzotto, Jacopo, Jesper Rüdiger, and Adrien Vigier, "Dynamic persuasion with outside information," *American Economic Journal: Microeconomics*, 2021, *13* (1), 179–94.

Blackwell, David, "Equivalent comparisons of experiments," *The annals of mathematical statistics*, 1953, pp. 265–272.

Bloedel, Alexander W and Ilya R Segal, "Persuasion with rational inattention," *Available at SSRN 3164033*, 2018.

Che, Yeon-Koo, Kyungmin Kim, and Konrad Mierendorff, "Keeping the listener engaged: a dynamic model of bayesian persuasion," *Journal of Political Economy*, 2023, *131* (7), 1797–1844.

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

Ely, Jeffrey C and Martin Szydlowski, "Moving the goalposts," *Journal of Political Economy*, 2020, *128* (2), 468–506.

Guo, Yingni and Eran Shmaya, "The value of multistage persuasion," Technical Report, working paper 2018.

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

Lehrer, Ehud and Dimitry Shaiderman, "Markovian Persuasion with Stochastic Revelations," 2022.

Ni, Bonan, Weiran Shen, and Pingzhong Tang, "Sequential Persuasion Using Limited Experiments," *arXiv preprint arXiv:2303.10619*, 2023.

Orlov, Dmitry, Andrzej Skrzypacz, and Pavel Zryumov, "Persuading the principal to wait," *Journal of Political Economy*, 2020, *128* (7), 2542–2578.

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

Senkov, Maxim, "Setting Interim Deadlines to Persuade," *arXiv preprint arXiv:2210.08294*, 2022.

Su, Shih-Tang, Vijay G Subramanian, and Grant Schoenebeck, "Bayesian Persuasion in Sequential Trials," *arXiv preprint arXiv:2110.09594*, 2022.

Appendix of Receiver Inattention and Persuading to Be Persuaded

A.1 Proof and Discussion of Lemma 3 and Proposition 1

Proof. Suppose that the optimal persuasion value in the static game is $1 - \rho(p^*_{q_1}, q_1)$ and there is a two-stage persuasion strategy where the sender chooses (ψ, q_2) for her first attempt and $(p^{*b}_{q_2}, p^{*g}_{q_2})$ for the final attempt if her first attempt fails, leading to the prior belief $q_2 < q_1$. For this two-stage persuasion strategy to attain the same value $1 - \rho(p^*_{q_1}, q_1)$ in expectation, it must hold that:

$$\rho(p^*_{q_1}, q_1) = \rho(p^*_{q_2}, q_2)\frac{\psi - q_1}{\psi - q_2}.$$

Solving this equation for ψ yields the argument (1) in Lemma 3. Here, with the predetermined parameter q_1, and the function $\mathcal{M}(\cdot)$, ψ is a function of q_2.

For a given $q_2 \in (0, q_1)$, in a optimal two-stage persuasion strategy, the sender chooses $(q_2, y(q_2, q_1))$ for the first, and $(p^{*b}_{q_2}, p^{*g}_{q_2})$ for the potential second attempt. The second attempt, if made, will generate an expected persuasion value of $1 - \rho(p^*_{q_2}, q_2)$. To match the overall persuasion value of this two-stage persuasion design with a static persuasion strategy, represented by $(\phi, y(q_2, q_1))$, the following condition must be met:

$$\frac{y(q_2, q_1) - q_1}{y(q_2, q_1) - \phi} = \rho(p^*_{q_2}, q_2)\frac{y(q_2, q_1) - q_1}{y(q_2, q_1) - q_2}.$$

Solving this equation for ϕ yields the argument (2) in Lemma 3. Given the predetermined parameter q_1, and the function $\mathcal{M}(\cdot)$, ϕ is a function of q_2. □

This benchmark, particularly $\psi(q_2)$, provides a convenient reference for comparing any two-stage persuasion strategy to the optimal persuasion strategy. It directly derives Proposition 1.

Proof. Note that in the design of the first persuasion attempt, $y(q_2, q_1)$ is the minimum feasible p^g_1, given that q_2 is chosen as p^b_1. Since the feasible set $\{p_1 | \mathcal{M}(p_1, q_1) \geq 0\}$ is compact, Lemma 1 implies that if $y(q_2, q_1) < \psi(q_2)$ for any $q_2 \in (0, q_1)$, then there exists a $\tilde{p}^g_1$ that is smaller than $\psi(\tilde{q}_2)$ when the given $\tilde{q}_2$ is chosen in the information structure as p_1^b. Given that $q_1 > q_2$, the derivative of $\frac{\psi - q_1}{\psi - q_2}$ with respect to ψ is positive. Therefore, choosing $(\tilde{q}_2, y(\tilde{q}_2, q_1))$ for the first attempt and $(p^{*b}_{\tilde{q}_2}, p^{*g}_{\tilde{q}_2})$ for the potential second attempt can make this sequential strategy outperform the optimal static persuasion strategy. When $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} \neq \emptyset$, $r(q_2) = \psi(q_2)$ can be an indicator of whether sequential persuasion strategy can outperform the optimal static strategy for the given q_2.

For certain values of q_2 where $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} = \emptyset$, the persuasion game remains static: when q_2 is chosen and the persuasion fails, the receiver will not pay attention to any further persuasion. Therefore, this strategy is always inferior to the one that optimizes the static persuasion game. In this case, let $r(q_2) = \frac{q_1 - \rho(p^*_{q_1}, q_1) q_2}{1 - \rho(p^*_{q_1}, q_1)}$. By definition that $\rho(p^*_{q_1}, q_1) = \frac{p^{*g}_{q_1} - q_1}{p^{*g}_{q_1} - p^{*b}_{q_1}}$, p_1^g is never smaller than $\frac{q_1 - \rho(p^*_{q_1}, q_1) q_2}{1 - \rho(p^*_{q_1}, q_1)}$. This definition ensures that $p_1^g \geq r(q_2)$ whenever $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} = \emptyset$, with equality only when $q_2 = p^{*b}_{q_1}$. As a result, we have $r(q_2)$ defined in Proposition 1 indicating that: for any $q_2 \in (0, q_1)$, if $y(q_2, q_1)$ is less than, equal to, or greater than $r(q_2)$, the optimal two-stage strategy conditional on $p_1^{*b} = q_2$ generates the overall persuasion value that is, respectively, greater than, equal to, and smaller than the optimal static persuasion value. □

By definition, the objective to minimize $\frac{p^g - q_t}{p^g - p^b}$ is quasi-convex and continuous in both $p_t = (p_t^b, p_t^g)$ and q_t. According to Lemma 1, the feasible set $\{p_t | \mathcal{M}(p_t, q_t) \geq 0\}$ is compact and convex, and the motivation function $\mathcal{M}(p_t, q_t)$ is continuous in q_t. Therefore, $\rho(p^*_{q_t}, q_t)$, as the optimized objective, is continuous in q_t. Let $\hat{q}_2$ be defined such that $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} = \emptyset$ if and only if $q_2 \leq \hat{q}_2$. Under this definition, $r(q_2)$ is continuous in $q_2 \in (0, \hat{q}_2] \bigcup (\hat{q}_2, q_1)$. If $\rho(p^*_{q_2}, q_2) \to_+ 0$ when $q_2 \to_+ \hat{q}_2$, then $r(q_2)$ is continuous for all $q_2 \in (0, q_1)$. In general, $r(q_2)$ is a continuous curve confined within the space, $\{(q_2, p_t^g) | \frac{q_1 - \rho(p^*_{q_1}, q_1) q_2}{1 - \rho(p^*_{q_1}, q_1)} \leq p_t^g \leq 1, 0 \leq q_2 \leq q_1\}$. Applying L'Hôpital's rule to $\psi(q_2)$ as $q_2 \to_- q_1$, we have $r(q_2) \to_+ q_1 - \frac{\rho(p^*_{q_1}, q_1)}{\rho'_{q_1}}$, where $\rho'_{q_1} < 0$ is the derivative of $\rho(p^*_{q_1}, q_1)$ with respect to q_1. Nevertheless, a more detailed characterization of $r(q_2)$, including its positioning and shape, relies on the specification of the motivation function.

A.2 Proof of Proposition 2

Proof. Based on Lemma 3 and Proposition 1, we establish the proof of Proposition 2 by analyzing $\psi(q_2)$ and $\phi(q_2)$.

When the information structure $(0, \gamma)$ is available for all $q_t \in (0, q_1]$ in a single persuasion attempt, the sender optimally chooses $(0, \gamma)$ for both the static persuasion attempt and the final attempt in a two-stage strategy. According to Lemma 3, this results in $\psi(q_2) \equiv \gamma$. Since $y(q_2, q_1) \geq \gamma$, by the structure of the persuasion problem central to this research, two-stage persuasion approach will never outperform the static one.

If $\mathcal{M}(p_t, q_t) \geq 0$ for all p_t and q_t, then no matter what q_2 is chosen for the first attempt, the sender will choose $p_2^* = p_{q_2}^* = (0, \gamma)$ for the final attempt, which generates $\rho(p_{q_2}^*, q_2) = \frac{\gamma - q_2}{\gamma}$. According to (2) of Lemma 3, we have

$$\phi(q_2) = q_2 \frac{\gamma - y(q_2, q_1)}{\gamma - q_2}.$$

Because $\mathcal{M}(p_t, q_t) \geq 0$ implies $y(q_2, q_1) \equiv \gamma$, it follows that $\phi(q_2) \equiv 0$. Therefore, no matter which q_2 is chosen, the optimal two-stage persuasion strategy is always equivalent to the optimal frictionless static strategy, represented by the information structure $(0, \gamma)$. Under this condition, since $(0, \gamma)$ is feasible in the static game, the optimal two-stage approaches, conditional on different q_2, are always equivalent to the optimal static approach in Bayesian persuasion. □

A.3 Proof of Proposition 3

Proof. Given that the motivation function has a Strictly Disappointment-Penalizing (SDP) feature, for any given $q_2 \in (0, q_1)$, we have:

$$\rho(p_{q_2}^*, q_2) \geq \max_{q > q_2} \rho(p_q^*, q),$$

indicating that $\rho(p^*_{q_2}, q_2)$ increases monotonically as q_2 decreases. Therefore, $\rho(p^*_{q_2}, q_2)$ is greater or equal to $\rho(p^*_{q_1}, q_2)$. Taking the derivative of $\psi(q_2)$ with respect to $\rho(p^*_{q_2}, q_2)$, the derivative is negative, provided $q_2 < q_1$. Setting $\rho(p^*_{q_2}, q_2) = \rho(p^*_{q_1}, q_2)$ in the expression of $\psi(q_2)$, we derive $\psi(q_2) = p^{*g}_{q_1}$. Because $\psi(q_2)$ is smaller when $\rho(p^*_{q_2}, q_2)$ is larger, the condition $\rho(p^*_{q_2}, q_2) \geq \rho(p^*_{q_1}, q_2)$ must yield $\psi(q_2) \leq p^{*g}_{q_1}$. Hence, as long as $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} \neq \emptyset$, it must be that $r(q_2) < p^{*g}_{q_1}$.

Notably, $p^g_1 < p^{*g}_{q_1}$ is not possible when $p^b_1 < p^{*b}_{q_1}$, otherwise $(p^{*b}_{q_1}, p^{*g}_{q_1})$ would not be an optimal static strategy. In fact, according to the definition of SDP, $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} = \emptyset$ holds for all $q_2 \leq p^{*b}_{q_1}$, because $\rho(p^*_{q_1}, p^{*b}_{q_1}) = 0$. If the sender chooses any information structure with $p^b_t \leq p^{*b}_{q_1}$, she will not have a second chance to persuade if her first attempt fails. Accordingly, choosing $p^b_1 > p^{*b}_{q_1}$ and $p^g_1 < p^{*g}_{q_1}$, if feasible, for the first persuasion attempt, is a necessary condition for a two-stage persuasion strategy to outperforms the optimal static persuasion strategy. To make this design possible, the chosen p^g_1 must also be smaller than $\psi(q_2)$, according to Proposition 1. □

A.4 Proof of Proposition 4

Proof. Suppose the sender's motivation has a Weakly Disappointment-Penalizing (WDP) feature, and generates $\psi(q_2)$ such that there exists a $y(q_2, q_1)$ that is smaller than $\psi(q_2)$. The sender's optimization problem becomes choosing an optimal q_2:

$$\min_{q_2} \frac{y(p^{*b}_{q_2}) - q_2}{y(p^{*b}_{q_2}) - p^{*b}_{q_2}} \cdot \frac{y(q_2, q_1) - q_1}{y(q_2, q_1) - q_2},$$

where $p^{*b}_{q_2}$, $y(p^{*b}_{q_2})$, and $y(q_2, q_1)$ all depend on q_2, conditional on the receiver's predetermined motivation function $\mathcal{M}(p_t, q_t)$ and the initial prior belief q_1.

The F.O.C. of this problem is shown as (5) in the main body. The component Λ in it is expressed as:

$$\Lambda = \frac{(q_1 - q_2) y'_{q_2}(q_2, q_1) + y(q_2, q_1) - q_1}{(y(q_2, q_1) - q_2)^2},$$

which is the derivative of $\frac{y(q_2,q_1)-q_1}{y(q_2,q_1)-q_2}$ with respect to q_2, where $y'_{q_2}(q_2, q_1)$ denotes the derivative of $y(q_2, q_1)$ with respect to q_2. According to this expression of Λ, in a static persuasion problem, an interior solution is possible only when $y'_{q_2}(q_2, q_1) < 0$ at $q_2 = p_1^{*b}$.

Given that the second-order condition is satisfied, $\Lambda \leq 0$ for $q_2 \leq p_{q_1}^{*b}$ and $\lambda \geq 0$ for $q_2 \geq p_{q_1}^{*b}$. Particularly, when Λ is continuous at $q_2 = p_{q_1}^{*b}$, then we have $\lambda = 0$ at this point. However, when the disappointment penalty is severe enough, which is represented by sufficiently small $y'_{q_2}(p_{q_2}^b, q_2) < 0$, the second term in (5) will be negative and sufficiently small to make the entire left-hand-side of (5) negative at $q_2 = p_{q_1}^{*b}$. Therefore, increasing $p_1^b = q_2$ above $p_{q_1}^{*b}$ will make the objective smaller, which benefits the sender. □

It is worth noting that although this proposition and the proof requires a sufficiently large disappointment penalty, this requirement is relative to the magnitude of $y'_{q_2}(p_{q_2}^b, q_2)$. If Λ is not continuous at $q_2 = p_{q_1}^{*b}$ and is positive when $q_2 \to_+ p_{q_1}^{*b}$, then its sign and value depend on the value of $y'_{q_2}(p_{q_2}^b, q_2)$. If $|y'_{q_2}(p_{q_2}^b, q_2)|$ is sufficiently large, the value of $\Lambda > 0$ is sufficiently small, allowing a relatively smaller $|y'_{q_2}(p_{q_2}^b, q_2)|$ to achieve $p_1^{*b} > p_{q_1}^{*b}$. Additionally, a more negative $y'_{q_2}(p_{q_2}^b, q_2)$ will cause the F.O.C. to remain negative even when p_1^b is way greater than $p_{q_1}^{*b}$ for a given level of disappointment penalty. This makes p_1^{*b} way larger than $p_{q_1}^{*b}$, which amplifies the cost of not recognizing that a two-stage persuasion strategy could outperform the static design

A.5 Proof and Discussion of Proposition 5

Proof. In the product of a sequential chain $\prod_{t=1}^{\tau-1} \frac{y(q_{t+1},q_t)-q_t}{y(q_{t+1},q_t)-q_{t+1}}$, which determines the persuasion value, inserting an additional component, $\frac{y(q_{t+1},q')-q'}{y(q_{t+1},q')-q_{t+1}}$, which immediately follows $\frac{y(q_{t+1},q_t)-q_t}{y(q_{t+1},q_t)-q_{t+1}}$, leaves the subsequent components unchanged. However, to navigate to this prior belief $q' \in (q_t, q_{t+1})$ upon a prior failed attempt, it requires the design of the attempt $\frac{y(q_{t+1},q_t)-q_t}{y(q_{t+1},q_t)-q_{t+1}}$ to change to $\frac{y(q',q_t)-q_t}{y(q',q_t)-q'}$. This modification allows all attempts preceding it to remain unchanged. To ensure the equivalence between the new and original persuasion design, the following condition must hold:

$$\frac{y(q_{t+1},q_t)+\delta_{t+1}-q'}{y(q_{t+1},q_t)-\delta_t-q'}\cdot\frac{y(q_{t+1},q_t)-\delta_t-q_t}{y(q_{t+1},q_t)+\delta_{t+1}-q_{t+1}}=\frac{y(q_{t+1},q_t)-q_t}{y(q_{t+1},q_t)-q_{t+1}},$$

where $y(q_{t+1},q_t)+\delta_{t+1}=y(q_{t+1},q')$ and $y(q_{t+1},q_t)-\delta_t=y(q',q_t)$.

Solving the above equation for δ_t produces Proposition 5, where $\rho(q_{t+1},q_t)=\frac{y(q_{t+1},q_t)-q_t}{y(q_{t+1},q_t)-q_{t+1}}$. Notably, when $\delta_{t+1}=0$, the proposition requires only $\delta_t=0$. □

In this proposition, the weak disappointment penalty implies $\delta_{t+1}>0$, which requires $\delta_t>0$ and that $y(q_{t+1},q_t)$ decreases with q_{t+1}. Without adjusting the design following the additional attempt, $p'=(q_{t+1},y(q_{t+1},q'))$, WDP is equivalent to SDP. To see this, suppose the inserted attempt with prior belief q' is the final attempt. When SDP is assumed, we have $\rho(p^*_{q'},q')\leq\rho(p^*_{q_t},q')$, implying that $p^*_{q'}$ is equivalent only to $(p^{*b}_{q_t},p^g)$, where $p^g>y(p^{*b}_{q_t},q_t)$. This scenario is equivalent to that which leads to Proposition 5.

If we allow for redesign of the persuasion path after the attempt $p'=(q_{t+1},y(q_{t+1},q'))$ to optimize the subgame conditional on the prior belief q', the required δ_t for making an additional persuasion attempt beneficial may reduce, or even turn negative, which eliminates the need for $y(q_{t+1},q_t)$ to decrease with q_{t+1}. However, this adjustment becomes less flexible and thus less likely as the length of the persuasion path following $p'=(q_{t+1},y(q_{t+1},q'))$ increases. For example, when the t^{th} attempt is the final attempt in the original design, a sufficiently small disappointment penalty for any $q<p^{*b}_{q_t}$ may be sufficient to remove the requirement that $y(q_{t+1},q_t)$ decrease with q_{t+1} for the additional (final) attempt to be beneficial. In contrast, when the path is longer, this disappointment penalty requirement extends beyond the t^{th} attempt in the original design. In other words, multiple weak disappointment-penalties can accumulate to create a more severe penalty effect over time.

A.6 Proof of Lemma 4 and Proposition 6

Proof. The sender's intertemporal persuasion problem over T periods is formalized as a Bellman equation as (6) in the main body. Taking the derivative of this equation with respect to q_{t+1} yields the F.O.C, or Euler condition, for the intertemporal problem:

$$\frac{y'_{t,q_{t+1}}}{p_t^{*g}-q_t}-\frac{y'_{t,q_{t+1}}-1}{p_t^{*g}-q_{t+1}}+\frac{y'_{t+1,q_{t+1}}-1}{p_{t+1}^{*g}-q_{t+1}}-\frac{y'_{t+1,q_{t+1}}}{p_{t+1}^{*g}-q_{t+2}}=0, \tag{A.1}$$

Rearrangements of (A.1) above produce Lemma 4 and Proposition 6.

Grouping terms according to $y'_{t,q_{t+1}}$ and $y'_{t+1,q_{t+1}}-1$, and leveraging the fact that $q_{t+1}=p_t^{*b}$ yields (7) in the Lemma 4.

When $y'_{t,q_{t+1}}<0$ and $y'_{t+1,q_{t+1}}<1$, we have $\frac{1}{p_{t+1}^{*g}-p_{t+1}^{*b}}<\frac{1}{p_t^{*g}-p_t^{*b}}$ according to (7), given that $q_t>q_{t+1}>q_{t+2}$. This implies $p_{t+1}^{*g}-p_{t+1}^{*b}>p_t^{*g}-p_t^{*b}$, which establishes the first argument in Proposition 6.

Furthermore, by grouping terms according to $y'_{t,q_{t+1}}$ and $y'_{t+1,q_{t+1}}$, condition (A.1) can be rearranged as follows:

$$\frac{1}{p_{t+1}^{*g}-q_{t+1}}=\frac{1}{p_t^{*g}-q_{t+1}}+y'_{t,q_{t+1}}\frac{q_t-q_{t+1}}{(p_t^{*g}-q_t)(p_t^{*g}-q_{t+1})}+y'_{t+1,q_{t+1}}\frac{q_{t+1}-q_{t+2}}{(p_{t+1}^{*g}-q_{t+1})(p_{t+1}^{*g}-q_{t+2})}.$$

According to this condition, when $y'_{t,q_{t+1}}<0$ and $y'_{t+1,q_{t+1}}<0$, we have $\frac{1}{p_{t+1}^{*g}-q_{t+1}}<\frac{1}{p_t^{*g}-q_{t+1}}$. This implies $p_{t+1}^{*g}>p_t^{*g}$, establishing the second argument in Proposition 6. □

A.7 Proof of Lemma 5 and Proposition 7

Proof. As a continued discussion of Proposition 5, Lemma 5 Suggests that the additional attempt $(q',y(q',q_t))$ should not replace $(q_{t+1},y(q_{t+1},q_t))$ and should instead be appended to it. This requires the following condition:

$$\frac{y(q_{t+2},q_{t+1})-q_{t+1}}{y(q_{t+2},q_{t+1})-q_{t+2}}\cdot\frac{y(q_{t+1},q_t)+\delta_{t+1}-q'}{y(q_{t+1},q_t)+\delta_{t+1}-q_{t+1}}<\frac{y(q_{t+2},q_{t+1})-\delta_{t+2}-q'}{y(q_{t+2},q_{t+1})-\delta_{t+2}-q_{t+2}}, \tag{A.2}$$

where $y(q_{t+2},q_{t+1})-\delta_{t+2}=y(q_{t+2},q')$ due to the disappointment penalty, and $q'>q_{t+1}$.

The terms $\delta_{t+2}>0$ and $\delta_{t+1}>0$ measures the disappointment penalty. Given that $q'>q_{t+1}>q_{t+2}$, when disappointment penalty decreases, the second term of the left-hand-side decreases, and the right-hand side increases. Therefore, the condition (A.2) is more likely to hold when the disappointment penalty is smaller. Particularly, there is no disappointment penalty (i.e., $\delta_{t+2} = \delta_{t+1} = 0$), the right-hand-side of (A.2) researches its maximum, reducing (A.2) to:

$$\frac{y(q_{t+1}, q_t) - q'}{y(q_{t+2}, q_{t+1}) - q'} < \frac{y(q_{t+1}, q_t) - q_{t+1}}{y(q_{t+2}, q_{t+1}) - q_{t+1}} \tag{A.3}$$

The left-hand side of the condition (A.3) decreases as q' increases. Additionally, it equals the right-hand side in (A.3) when $q' = q_{t+1}$. Given that $q \in (q_{t+1}, q_t)$, condition (A.3) always holds. Because in (A.2), left-hand side and right-hand side are continuous in δ_{t+1} and δ_{t+2}, respectively, (A.2) is always true when disappointment penalty is sufficiently small. This completes the proof of Lemma 5. □

Proposition 7 is a direct conclusion of Proposition 5 and Lemma 5.

Proof. When the disappointment penalty approaches 0, δ_{t+1} also approaches 0 by definition. Therefore, as long as an information structure is not mutually dominated by any information structure in a given path, $\delta_t > 0$ and Proposition 5 is satisfied. This makes it beneficial to include this information structure in the persuasion path. According to Lemma 5, this inclusion does not replace any existing information structure in the path. Since Proposition 5 provides a sufficient condition for including additional persuasion stages, if any non-mutually-dominated information structure is omitted in the persuasion path, the design cannot be optimal, given that the attempt limit is not binding. As a result, an optimal persuasion design must incorporate all non-mutually-dominated information structures.

Specifically, when there is no disappointment penalty, $y(p^b, q_t) \equiv y(p^b, q_1)$ for all $q_t \in (0, q_1]$. In this case, $\big(p^b_{min}, y(p^b_{min}, q_1)\big)$ should not be mutually dominated by any feasible information structure and should be part of an optimal design. This information structure, $\big(p^b_{min}, y(p^b_{min}, q_1)\big)$ should be designed as a final attempt in period τ because $q_{t+1} < q_t$ and no smaller p^b exists once q_t reaches p^b_{min}. Given this optimal subgame, we can construct an optimal persuasion path using backward induction. For the $\tau - 1^{th}$ attempt, p^b should be chosen as $\min\{p^b | y(p^b, q_1) < y(p^b_{min}, q_1)\}$ so that this information structure and $\big(p^b_{min}, y(p^b_{min}, q_1)\big)$ are not mutually dominated. p^g_{min} should be chosen for the information structure in the first attempt. To see this, for any $\prod \frac{y(q_{t+1},q_t)-q_t}{y(q_{t+1},q_t)-q_{t+1}}$, where p^g_{min} is not included, introducing $(y^{-1}(p^g_{min}, q_1), p^g_{min})$, where $y^{-1}(\cdot)$ is the inverse function, in place of the information structures that are mutually dominated by $(y^{-1}(p^g_{min}, q_1), p^g_{min})$ benefits the sender. After this adjustment, $(y^{-1}(p^g_{min}, q_1), p^g_{min})$ should be the first persuasion attempt considering Bayes plausibility. This completes the proof of Proposition 7. □