Receiver Inattention and Persuading to Be Persuaded

Mengqi Zhang*

June 26, 2025

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, Jeffrey Ely, 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.

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 such as 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 not be able 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 a 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, almost each attempt was a "elevator pitch", with only limited attention from Sculley. Since persuasion rarely succeeds in the first attempt, it is important to make the receiver persuadable on the following attempts. 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 the persuasion that concludes within a single attempt, Alice needs 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 persuasion strategy versus sequential persuasion strategy generally differs, particularly in the sender's first attempt. In a sequential design, the first attempt serves as an opening, while 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 so that a bad signal substantiates the bad state, the sender maximizes the probability of the “good” signal, 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 patient and make him open for 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. Although each version provides a complete analysis of the same topic, the time required to finish the pitch, such as 1 minute, 5 minutes, or 10 minutes, 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 provides a possible microeconomic foundation to explain the observation of Che et al. (2023): each persuasion attempt may not necessarily result in a “jump” over the cutoff point in the receiver’s belief, but it best preserves the receiver’s patience while serving to gradually persuade the receiver. This approach allows multiple attempts and technically expands the receiver’s participation constraint, 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 sequential persuasion approach is more effective than the static one under the two-stage framework (Proposition 1, Proposition 2). Our benchmark analysis 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 the optimal static one (Corollary 1). Naturally, the receiver may become disappointed after a failed attempt, reducing the feasible set of information structures for the sender's future persuasion efforts. If the receiver's motivation is sensitive to previous failures, the sender must adopt a more conservative persuasion strategy in the initial attempt (Proposition 4); otherwise, she risks a discount in effectiveness or even losing the opportunity for further persuasion (Proposition 3).

Generalizing these findings, we examine the conditions under which the sender benefits from extending the persuasion path beyond two stages, where she chooses a "piecemeal" information disclosure strategy (Proposition 5, Proposition 6). Based on this incremental disclosure feature of the sequential persuasion design, we show that if the receiver's motivation to participate does not deteriorate substantially after previous failures, the sender always exhausts all permitted persuasion attempts whenever she finds any sequential persuasion design outperforms the optimal static persuasion (Proposition 7). This result simplifies the sender's design problem and underscores that the two-stage analysis remains central even in more general multistage settings. It also provides a baseline from which we can predict the sender's practical strategy when conditions depart from the ideal.

To further assess to what extent a sequential approach can beat frictional constraints due to the receiver's inattention, we also allow the maximum number of persuasion attempts to approach infinity. While extending the persuasion process may improve effectiveness, we find that the overall efficiency boundary remains fundamentally constrained by the receiver's inattention. Therefore, sequential persuasion cannot fully recover the efficiency loss resulting from this friction (Proposition 8). An illustrative example underscores this finding and provides additional insight that the effectiveness of the sequential persuasion approach as a remedy for receiver inattention diminishes as inattention becomes more severe (Corollary 3).

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,[1] 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 established 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" in 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

[1] 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.

when the attempt limit is sufficiently large to be non-binding.

This special intertemporal framework provides theoretical insight into decision making in information design. By characterizing the sender's responses to frictional constraints, it explains the sender's conservative persuasion behaviors, such as "piecemeal information disclosure", which could 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 improve 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, which are often unobservable, to the observable decisions. With this theoretical foundation, calibrating parameters in the information cost as functions 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 arbitrary feasible sets of information structures as frictional constraints, along with indifference curves representing success probabilities based on prior beliefs (Figure 2). By analyzing the boundary shapes of the feasible set and how they shift over time with changing prior belief, this system enables direct comparisons of the effectiveness of sequential persuasion strategies and helps characterize constraint properties that influence sender decisions. This analytical approach not only enhances the tractability of our model, but also intuitively illustrates the mechanism through which sequentiality relaxes frictional constraints, demonstrating its power in persuasion games. Moreover, because the framework accommodates arbitrary feasible sets of information structures, our analysis remains general rather than being restricted to specific forms of information cost functions. With these advantages, this system is readily applicable to a broad range of scenarios, extending beyond the scope of this study while remaining adaptable within a similar theoretical framework to provide both tractable analysis and 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.[2] 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 accordance with the 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 Model

To illustrate how a sequential persuasion strategy can improve persuasion effectiveness when information structures are subject to frictional constraints, we present a numerical toy model.

[2]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.

In the elevator pitch example, assume that 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 a 0 value. Although Alice understands her business idea, she is inexperienced and uncertain of its market viability. In this case, the information is symmetric between Bob and Alice at the beginning. They both believe that 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 represented in table 1: 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 the signal b appears, he rejects the request. Under this information structure, Alice has a 5/6 chance of getting 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 pitch outside 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 failure in the first attempt also eliminates any further opportunities to succeed.

In 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 drops 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 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 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. Keeping 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 2 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 model, despite its stylized feasible set, provides intuitive answers to the main questions of this study. 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 frictional constraints, as represented by A in this example. This remains valid even when a previous failure in persuasion makes the receiver less patient, thus 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 attempt at persuasion 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 in which 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 value of the product is 1 when $\omega = h$ and 0 otherwise. The sender's payoff $u = \gamma\alpha$, 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 publicly known information structures $\pi : \Omega \to \Delta\Theta$, which map states to a full support distribution over possible signals. Upon receiving a signal that is drawn publicly according to the information structure, 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 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 commits to using this information structure to generate the signal that is drawn publicly, 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 to pay attention to the persuasion in period t as a time-invariant function of the prior belief q_t and the information structure π_t, as follows:

$$\mathcal{M}(q_t, \pi_t) = \sum_{\theta \in \Theta} \max\left\{0,\ \pi_t(\theta|h)q_t - \gamma\left[\pi_t(\theta|h)q_t + \pi_t(\theta|l)(1-q_t)\right]\right\} - c(q_t, \pi_t), \tag{1}$$

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 1, 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 the function 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 nonnegative and continuous in q_t, deferring further specification until the next section, where we discuss the model simplification. The specification based on the simplification 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. We assume that the sender has to pay a negligible but positive cost every time the experiment is conducted. Thus, the sender stops persuasion if there is no information structure that both meets the receiver's participation constraint ($\mathcal{M}(q_t, \pi_t) \geq 0$) and possibly changes his belief, or if the receiver's prior belief has already supported 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 1 below.

The timeline in the figure 1 indicates that the sender cannot predict the signal realization, and thus does not know ex ante when to 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

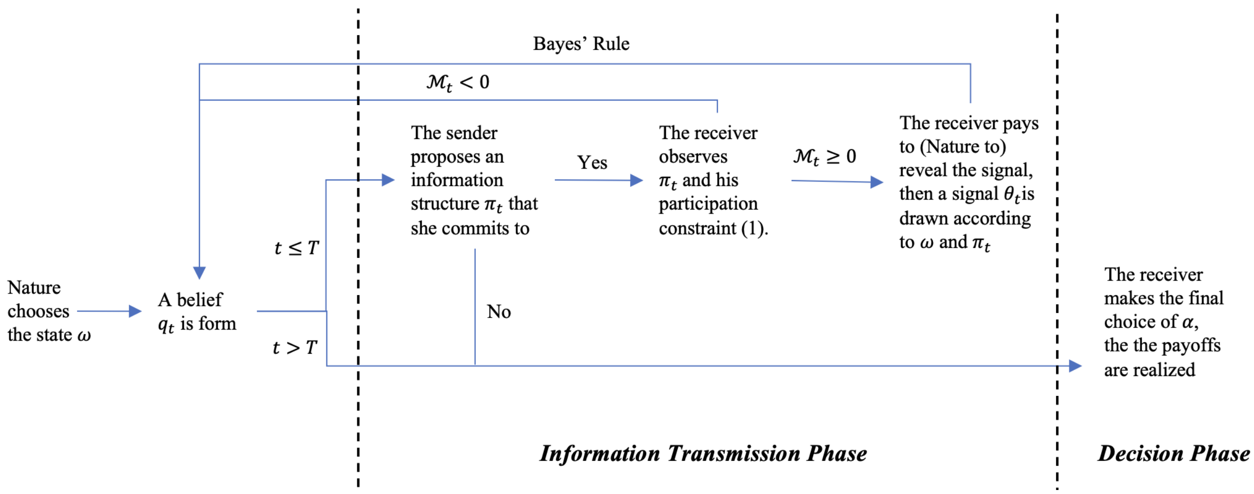


Figure 1: Timeline of an Endogenous Sequential Bayesian Persuasion Game

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 minimize the overall probability of no signal recommending "act". The sender's objective throughout the information transmission phase is therefore structured as follows:

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

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 the Bayes plausibility and the frictional constraint $\mathcal{M}(q_t, \pi_t) \geq 0$ due to the receiver's inattention. In addition, 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, which sets 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 participation constraint in each period t is binding, the receiver is indifferent between the different information structures proposed at different stages. Accepting any proposal that satisfies the constraint is 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 establish 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 "act" or "not act". Let $g, b \in \Theta$ denote the good and bad signals recommending "act" and "not act", respectively. Accordingly, p_t^g and p_t^b denote the posterior beliefs about $\omega = h$ in period t. Then a non-dominated Bayes-plausible strategy, $p_t = (p_t^b, p_t^g)$, requires $p_t^g \geq \gamma$ 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 upper semicontinuous in p_t for all $p_t \in [0, q_t] \times [\gamma, 1]$, then there exists an optimal strategy $p_t^* \in [0, q_t] \times [\gamma, 1]$ for the problem 3.*

The upper semicontinuity 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. *When the information cost $c(p_t, q_t) > 0$ for all p_t and q_t, the receiver pays attention to the persuasion in the period t only if $p_t \in [0, q_t] \times (\gamma, 1]$.*

According to Lemma 2, even before the final persuasion attempt, the sender should adopt an experiment that recommends the receiver to either "act" or "not act" if the receiver must always incur a positive cost to reveal the signal. In this case, the receiver's participation constraint is satisfied only when $p_t \in [0, q_t) \times (\gamma, 1]$, which ensures a positive information value. Whenever the sender makes an take-it-or-leave-it persuasion offer at any stage of a finite dynamic game, the receiver will accept it. Therefore, even if a posterior may lead to a subgame with a positive expected value, the sender cannot commit to sharing this additional value with the receiver in the future. With this anticipation, the receiver has zero patience. He requires the persuasion to possibly yield an immediate reward to capture his attention. Given this, the sender's goal of influencing the receiver's decision rather than only adjusting his belief for future persuasion applies not only to the final persuasion attempt, but to any node on the designed 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 that "bad" signals appear 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.

In addition to 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 some 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 subsequent persuasion attempts of the sender 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 $\mathcal{M}(p_t, q_t) \geq 0$ conditional on prior belief q_t and choice of p_t^b. Furthermore, 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 following definition 1 specifies 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. The definition of disappointment-penalizing forms the axiom underlying our subsequent analysis.

Axiom 1. *The receiver's motivation function satisfies Lemmas 1 and 2, and is strictly quasi-concave. Additionally, it is weakly disappointment-penalizing.*

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 creates 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 belief "sweet spots" and the sender's ad hoc motivation to prolong the game, which not only aligns the model more closely with the real world, but also ensures that the research question remains economically 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: $q_{t+1} < q_t$. In 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 position q_{t+1} and to the value function at position $y(q_{t+1}, q_t)$. The 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 second (final) persuasion attempts are colored 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 Sequential 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 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.

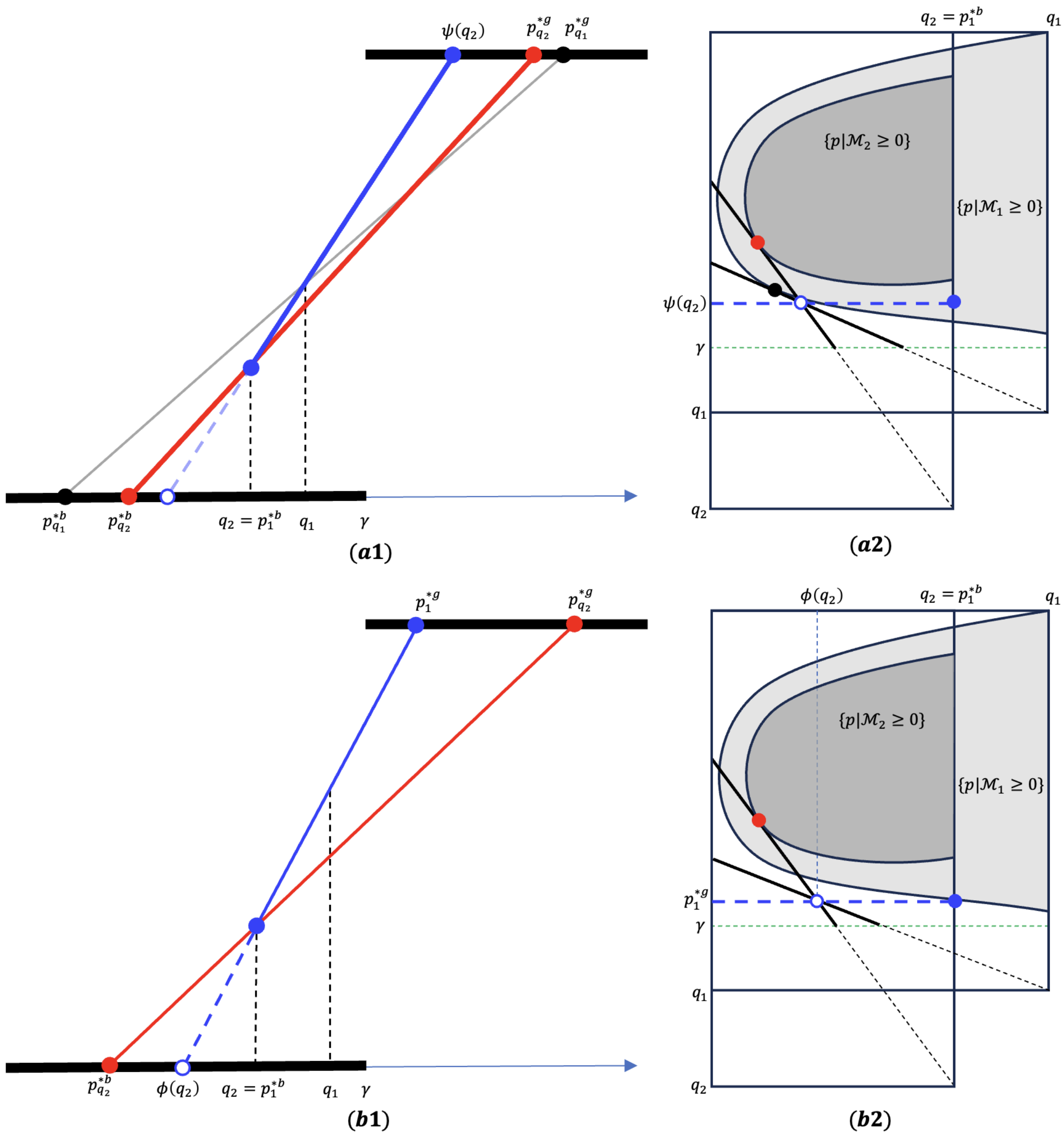


Figure 2: Equivalence Between Strategies with Feasible Sets

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 $\frac{q_t - p_t^b}{p_t^g - p_t^b}$. In the graph, this value is represented as the slope of the linear indifference curve that connects $p_t = (p_t^b, p_t^g)$ 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 be transformed into equivalent comparisons of 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 $q_2 \in [0, q_1)$, 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 $q_2 \in [0, q_1)$, 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 a two-stage strategy to be equivalent to the optimal static strategy $(p_{q_1}^{*b}, p_{q_1}^{*g})$ and also to have q_2 chosen as the prior belief that defines the second-stage persuasion problem, this line segment must attain the persuasion value $1 - \rho(p_{q_1}^*, q_1)$ at position q_1 and $1 - \rho(p_{q_2}^*, q_2)$ at position q_1. 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, it is sufficient to compare feasible first-stage information structures $(q_2, y(q_2, q_1))$ with first-stage information structures in these benchmarks $(q_2, \psi(q_2))$ conditional on the same q_2.

From a different perspective, the second half of Lemma 3 translates optimal sequential strategies, conditional on different choices of q_2, into equivalent static strategies. To find such static strategies, the blue line segment representing the first-stage strategy is extended so that its both ends connect to the value function in the panel $(b1)$. In panel $(b2)$, this equivalent static strategy is identified by projecting the (blue) 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 participation constraint 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 approach outperforms static approach in the sender's persuasion attempt 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, the sender's 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 success rate had she needed to persuade for the second time, $1 - \rho(p^*_{q_2}, q_2)$, and her probability of success in 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 of 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 success in the first attempt increases without reducing the persuasion value of a possible second attempt. As a result, the sequential strategy outperforms the optimal static strategy.

The proposition 1 simplifies the comparison between optimal static and sequential persuasion strategies by reducing it to a comparison between feasible first-stage strategies and the benchmark. This framework makes it convenient to determine whether the sequential approach is more effective than the static one in persuasion once the profile of a persuasion game is given. Within this framework, we can also answer the general question of what makes a sequential (two stages) approach outperform the static one. The answer to this question depends on how the properties of the functions $r(q_2)$ and $y(q_2, q_1)$ are shaped under different conditions.

5.2 Necessary and Sufficient Conditions for Endogenous Sequentiality

How $\mathcal{M}(p_t, q_t)$ varies with p_t and q_t shapes $y(q_2, q_1)$ and $r(q_2)$. For $y(q_2, q_1)$, this is straightforward. As the lower boundary of the feasible information structure set for a given prior belief q_1, its variation with q_2 reflects how $\mathcal{M}(p_t, q_t)$ changes in the information structure p_t. On the other hand, $\rho(p^*_{q_t}, q_t)$ determines $r(q_2) = \psi(q_2)$, provided that $r(q_2) > y(q_2, q_1)$ is possible. Given the same persuasion objective, an optimization outcome partially measures the constraint. Thus, how $\mathcal{M}(p_t, q_t)$ varies with the receiver's prior belief q_t is partially captured by $\rho(p^*_{q_t}, q_t)$ and, consequently, by $r(q_2)$. Since the disappointment-penalizing feature in Definition 1 relates to how $\mathcal{M}(p_t, q_t)$ varies with q_t, it also affects $r(q_2)$. Therefore, our discussion of the comparison between $y(q_2, q_1)$ and $r(q_2)$, which determines whether the sender is motivated to extend the persuasion game, should be based on the disappointment penalty.

Proposition 2. *Given Axiom 1, the sender chooses a two-stage persuasion over a static persuasion only if there exists* $q_2 \in [0, q_1)\backslash\{p^{*b}_{q_1}\}$ *such that* $\frac{\partial y(q_2,q_1)}{\partial q_2} < 0$*. This condition is also sufficient when the disappointment penalty is sufficiently small.*

Proof. Proof of Proposition 2. See Appendix.

Proposition 2 provides an answer to the question of when the sender is motivated to extend the static persuasion game to the sequential one under axiomatic conditions. This answer suggests a lower boundary of the feasible set of information structures with a downward-sloping portion. By Lemma 3 and Proposition 1, with a zero disappointment penalty, the most lenient case under Axiom 1, $r(q_2)$ is above $p^{*g}_{q_1}$ for any $q_2 \geq p^{*b}_{q_1}$. In this case, a lower boundary with a negatively sloped portion suffices to motivate the sender to adopt a sequential approach to persuade. This sufficiency holds as long as the disappointment penalty remains sufficiently small.

The necessity of Proposition 2 follows intuitively from its contrapositive. If the entire lower boundary of the feasible set is upward-sloping, the optimal static persuasion strategy must use the globally smallest p^b and p^g. In this case, $r(q_2) \leq p^{*g}_{q_1} \leq y(q_2, q_1)$ for all possible q_2, which contradicts the condition in Proposition 1. Therefore, the sequential approach never outperforms the static one. This implies that sequential persuasion never outperforms static persuasion, which establishes the contrapositive. An important corollary of this necessity is that the sequential approach never outperforms the static one without friction because the lower boundary of the feasible set defined only by Bayes plausibility is horizontal. This corollary is formalized as follows.

Corollary 1. *If 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. If the strategy $p = (0, \gamma)$ is always feasible within the static persuasion approach, then the sequential persuasion approach never outperforms it.*

Proof. Proof of Corollary 1. See Appendix.

Corollary 1 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 $t = 1$. Corollary 1 shows that the benefits and costs associated with this belief change perfectly offset each other, as reflected by $\phi(q_2) \equiv 0$, according to the statement 2) of Lemma 3.

When the persuasion game lacks a predetermined dynamic and neither information costs nor discounting is considered, the sender is indifferent between a static and sequential persuasion approach. Therefore, it is justifiable that previous conventional studies opt for a static framework without sacrificing generality. However, when frictional constraints emerge due to the receiver's inattention, a static framework becomes insufficient to fully capture the sender's persuasion strategy. This first argument of Lemma 3 highlights the need for a framework that accounts for the sender's choice of sequentiality in the persuasion game and clarifies the relationship between persuasion sequentiality and friction.

Second, even if the receiver does not pay attention to all Bayes-plausible persuasion strategies, as long as $p = (0, \gamma)$ remains a feasible information structure, it becomes the sender's optimal strategy in a static persuasion game. In this case, $r(q_2) \equiv \gamma$, it is always true that $r(q_2) = \gamma \leq y(q_2, q_1)$. 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. A more important implication of this corollary is that a frictional constraint, which restricts information structures beyond 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 as Proposition 2.

Typically, with the static optimal information structure, the persuasion game reaches efficiency, where any improvement of the sender's persuasion value comes at the cost of the receiver's information value. The condition indicated in Proposition 2 requires that when the information structure deviates from the static optimum, the loss of this efficiency is sufficiently small. It allows the sender to sacrifice as little of the current persuasiveness as possible to secure an additional persuasion opportunity that generates extra persuasiveness.

Panel (b2) in Figure 2 illustrates this underlying mechanism. The sequential approach allows the sender to de facto use a smaller p^b without changing p^g, represented by the horizontal projection to the left from the (blue) point denoting the first stage persuasion strategy. This introduces an alternative trade-off mechanism that can potentially replace the single-stage mechanism, where the trade-off occurs along the lower boundary of the feasible set of information structures. When the lower boundary between the auxiliary static strategy, represented by the intersection of indifference curves on the graph, and the first-stage strategy is upward-sloping, the trade-off along the constraint boundary dominates, making the sequential approach unnecessary. However, when this boundary is downward-sloping, positioning it above p_1^g when p^b is smaller than p_1^b, the trade-off mechanism in the sequential approach demonstrates its power. It allows the sender to achieve a de facto information structure beyond the constraint, attaining the overall persuasiveness that is otherwise impossible with only a static persuasion attempt.

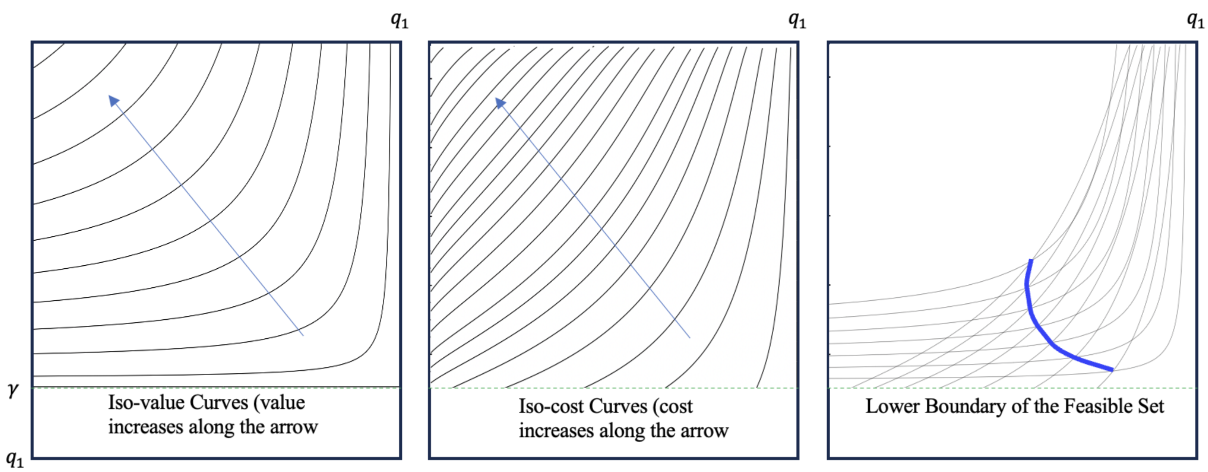


Figure 3: Information Cost Function and Boundary of the Feasible Set

The variation of the information cost with the informativeness of the information structure is necessary to create the condition indicated in Proposition 2 that favors the sequential approach. According to the first term of equation 1, switching to a different information structure while maintaining the same information value for the receiver requires p^b and p^g to shift in opposite directions. This is illustrated in the first panel of Figure 3, where all iso-value curves exhibit non-negative slopes. With a fixed information cost, one of these curves forms the lower boundary of the feasible set of information structures, or the participation constraint. However, if information cost varies with informativeness, a set of iso-cost curves emerges within the region where Bayes plausibility holds. In this case, the boundary of the participation constraint is determined by the intersection of the iso-value and iso-cost curves with the same value. For example, when the information cost is a function of mutual information, the iso-cost curves in the middle panel of Figure 3 establish the participation constraint boundary as the bolded (blue) curve in the last panel, which has a downward slope.

Mutual information is widely used as a determinant of information cost in related research. As shown in Figure 3, the shape and the gradient direction of its resulting iso-cost curves allow for the formation of a downward sloping constraint boundary of any desired shape, provided the gradient magnitude is appropriate. More broadly, information costs satisfying Blackwell's Theorem (1953) may exhibit similar properties. This suggests that the sender's motivation to extend the persuasion process may be a common feature across various persuasion games where the receiver's inattention is considered. Incorporating an endogenous sequential structure into these models could enhance the robustness of existing findings and provide new insights.

6 A Less Informative Opening Pitch

Although persuasion with multiple attempts may increase the overall chance of convincing the receiver compared to a single optimal attempt, it does not necessarily need to be formalized as a sequential persuasion problem. In some instances, information structures in different stages of an optimal sequential strategy may optimize these persuasion attempts as if they were independent problems. When this occurs, treating these attempts as repeated static persuasion problems is sufficient, making an intertemporal framework as well as the analysis in the previous subsection unnecessary. In this subsection, we discuss the sender's optimal strategy in persuasion with multiple attempts, which not only answers the question of how the sender extends the persuasion process, but also determines the importance of the question of when to extend it. Since the disappointment-penalizing feature shapes the sender's intertemporal trade-off, the discussion of the sender's optimal persuasion strategy requires conditioning on different levels of disappointment penalty.

Proposition 3. *If the motivation function is SDP, 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* $y(q_2, q_1) < r(q_2)$.

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. When 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 can choose a p_1^b above $p_{q_1}^{*b}$. Although this less aggressive strategy has a lower chance of convincing the receiver instantly, it renders a "bad" signal "less bad", preserving the possibility of a second persuasion attempt if the first fails.

The sender who faces the receiver's SDP participation constraint should be aware that if there is any opportunity to increase the effectiveness of persuasion with multiple attempts, she needs to realize it through optimizing an intertemporal problem, which requires her to modify conventional static strategy for a potential second persuasion attempt. If she fails to do so, she will completely miss out on this opportunity. Such a stark difference in outcomes may encourage the sender to pay closer attention to avoid such an oversight.

Failing to recognize that a persuasion game should be sequential in optimal has more subtle consequences when the motivation function is WDP but not SDP. A WDP motivation function may not eliminate the possibility of subsequent persuasion when the sender initially chooses $p^*_{q_1}$. In this case, sequential persuasion outperforms the optimal static strategy by allowing the sender to gain additional persuasion value without reducing persuasiveness on first attempt. However, the optimal static strategy may not be the best choice for the initial attempt in a two-stage sequential design. If the sender fails to recognize the advantage of sequential strategies and mistakenly chooses the strategy that is only optimal for the static approach, she risks missing the opportunity to implement the most effective persuasion design, even if a second attempt remains available when the receiver's prior belief shifts to $p^{*b}_{q_1}$.

Proposition 4. *Under Axiom 1, suppose that the disappointment penalty is zero. Define:*

$$\bar{s}(q_2) = \frac{(p^{*g}_{q_1} - q_1)q_2\hat{\rho} - (\rho^*_{p^*_{q_1}} - \rho^*_{q_1})q_1q_2 + (q_1\rho^*_{p^*_{q_1}} - p^{*b}_{q_1}\rho^*_{q_1})q_1}{(\hat{\rho} - \rho^*_{q_1})p^{*b}_{q_1} + (\rho^*_{q_1} - \rho^*_{p^*_{q_1}})q_2 + (\rho^*_{q_2} - \hat{\rho})q_1},$$

and

$$\underline{s}(q_2) = \max\left\{\frac{(p^{*g}_{q_1} - p^{*b}_{q_1})q_2\hat{\rho} - (p^{*g}_{q_1} - q_2)q_1}{(p^{*g}_{q_1} - p^{*b}_{q_1})\hat{\rho} - (p^{*g}_{q_1} - q_2)}, \frac{(1 - p^{*b}_{q_1})q_2\hat{\rho} - (1 - q_2)q_1\rho^*_{p^{*b}_{q_1}}}{(1 - p^{*b}_{q_1})\hat{\rho} - (1 - q_2)\rho^*_{p^{*b}_{q_1}}}\right\},$$

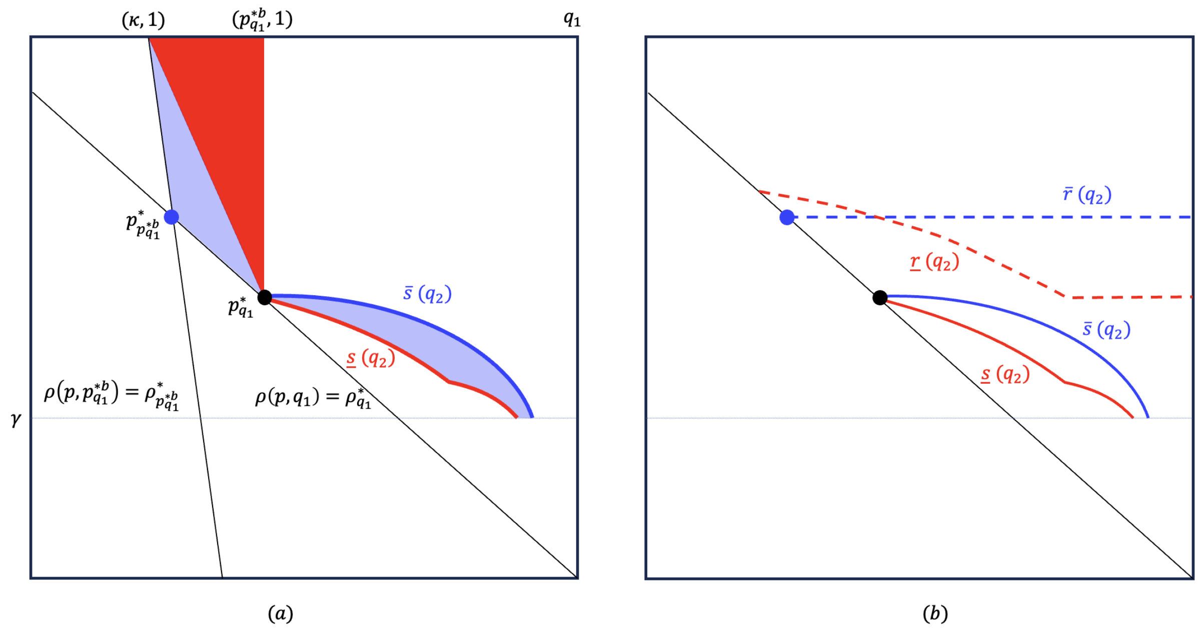
where p^*_q and $1 - \rho^*_q$ denote the optimal static persuasion strategy and value when the prior belief is q, and $1-\hat{\rho} = 1-\rho^*_{p^*_{q_1}}\rho^*_{q_1}$ represents the optimal overall persuasion value in the two-stge sequential persuasion approach conditional on the prior belief being q_1 and choosing $p^*_{q_1}$ for the first stage of persuasion. Assume that $\rho^*_{p^*_{q_1}} < 1$. The sender chooses $p^{*b}_1 > p^{*b}_{q_1}$ when there exist q_2 such that $y(q_2, q_1) < \underline{s}(q_2)$ and only when there exist q_2 such that $y(q_2, q_1) < \bar{s}(q_2)$.

Proof. Proof of Proposition 4. See Appendix.

To know which alternative sequential strategies can produce the same overall persuasiveness as in this circumstance, we can construct a benchmark, $s(q_2)$, to measure the necessary levels of p_1^g when different $q_2 = p_1^b$ are chosen, then based on which the second persuasion attempt is optimized, similarly to Proposition 1. When the static optimal persuasion value $\rho^*_{q_1}$ is also given, Proposition 4 outlines the upper bar $\bar{s}(q_2)$ and the lower bar $\underline{s}(q_2)$, of $s(q_2)$ in the case where the disappointment penalty is absent. Then, $y(q_2, q_1)$ being smaller than these two boundary benchmarks, respectively, becomes the necessary and sufficient condition that motivates the sender to choose a more conservative strategy as the opening of a sequential persuasion attempt.

With the assumption that $\rho^*_{p^*_{q_1}} < 1$ in Proposition 4, the sender retains the opportunity for subsequent persuasion even if she initially persuades as if she had only a static chance. To identify alternative sequential strategies that achieve the same overall persuasiveness in this setting, we define a benchmark $s(q_2)$, to determine the required levels of p_1^g given different choices of $q_2 = p_1^b$ and the optimal persuasion value in the subsequent attempt based on prior belief q_2. This is a similar approach to Proposition 1. When the static optimal persuasion value $\rho^*_{q_1}$ is also given, Proposition 4 establishes the upper and lower bounds, $\bar{s}(q_2)$ and $\underline{s}(q_2)$, for $s(q_2)$ in the absence of a disappointment penalty. The conditions $y(q_2, q_1) < \bar{s}(q_2)$ and $y(q_2, q_1) < \underline{s}(q_2)$ are then necessary and sufficient, respectively, for employing a different pitch in sequential persuasion than the optimal static strategy $p^*_{q_1}$. Compared to $p^*_{q_1}$, this optimal opening pitch features a larger p^b and a smaller p^g, making it less informative and more conservative.

Panel (a) of Figure 4 illustrates this concept based on two indifference curves for optimal static persuasion values, conditional on prior beliefs q_1 and $p^{*b}_{q_1}$. The shaded quadrilateral to the upper left of $p^*_{q_1}$ represents the region that contains the feasible constraint left of $p^{*b}_{q_1}$. If this constraint includes the point $p^* p^{*b}_{q_1}$, then for any prior belief $q_2 \in [p^{*b}_{q_1}, q_1]$, this point remains an



Note: $\kappa = 1 - \frac{1-p^{*b}_{q_1}}{\rho^*_{p^{*b}_{q_1}}}$

Figure 4: Possible Region of the Determination $s(q_2)$

optimal strategy for the second-stage persuasion attempt. Here, $\bar{s}(q_2)$ serves as a benchmark to determine whether a more conservative opening pitch is warranted. Alternatively, if the boundary of the feasible set lies entirely within the region $p^*_{q_1}(\kappa, 1)(p^{*b}_{q_1}, 1)$, then the optimal strategy for the second-stage persuasion is $p^*_{q_1}$ for $q_2 \in (\hat{\kappa}, q_1]$ and $(\kappa, 1)$ for $q_2 \in [p^{*b}_{q_1}, \hat{\kappa}]$, where $\hat{\kappa} = \frac{\kappa p^{*g}_{q_1} - p^{*b}_{q_1}}{\kappa - 1 + p^{*g}_{q_1} - p^{*b}_{q_1}}$. In this case, $\underline{s}(q_2)$ becomes the relevant benchmark. For other situations where the boundary lies entirely within the shaded quadrilateral, $s(q_2)$ falls within the shaded area between $\bar{s}(q_2)$ and $\underline{s}(q_2)$.

Compared to $r(q_2)$, which determines whether a static or sequential approach is preferable, $s(q_2)$ requires a lower feasible set boundary to justify a more conservative opening pitch than the optimal static persuasion. Panel (b) of Figure 4 illustrates this distinction. Given a zero disappointment penalty, Proposition 2 implies that the existence of (p^b, p^g) in the lower left region of $p^*_{q_1}$ is sufficient to favor sequential over static persuasion. However, Figure 4 shows that this condition alone is not sufficient to also motivate the sender to modify the persuasion strategy in the first stage. If the feasible constraint supports $\rho(q) = p^*_q < 1$ for both prior beliefs q_1 and $q = p^{*b}_{q_1}$, then a sequential strategy with a conservative opening pitch must outperform both the optimal static strategy and the sequential design that begins with $p^*_{q_1}$ in the first attempt. This results in $s(q_2) < r(q_2)$.

Corollary 2. *Under Axiom 1, suppose that there are different motivation functions $\mathcal{M}_1(p_t, q_t)$ and $\mathcal{M}_2(p_t, q_t)$ such that $\mathcal{M}_1(p_t, q_t) \equiv \mathcal{M}_2(p_t, q_t)$ when $q_t = q_1$, $\{p_t | \mathcal{M}_1(p_t, q_t) \geq 0\} \subsetneq \{p_t | \mathcal{M}_2(p_t, q_t) \geq 0\}$ when $q_t < q_1$, and only $\mathcal{M}_1(p_t, q_t)$ is SDP. If $\mathcal{M}_1(p_t, q_t)$ motivates the sender to choose persuade sequentially, she optimally chooses $p^{*b}_1 > p^{*b}_{q_1}$ when the receiver's motivation function is $\mathcal{M}_2(p_t, q_t)$ with sufficiently large disappointment penalty.*

According to Corollary 2, even if a sequential strategy $(p_1, p^*_{p_1})$ can survive the SDP feature to outperform the optimal static strategy, when the motivation function instead does not impose a disappointment penalty, it may still be inferior to a sequential strategy where $p^* q_1$ is chosen in the first attempt. In Figure 4, when the motivation function is SDP, the relevant benchmark $r_{SDP}(q_2)$ is represented by an upward-sloping curve approaching $p^{*g}_{q_1}$ as q_2 approaches q_1. Although $r_{SDP}(q_2)$ lies below $\bar{r}(q_2)$ and $\underline{r}(q_2)$, it can cross the region between $\bar{s}(q_2)$ and $\underline{s}(q_2)$, meaning that a strategy p within this region may fall below $r_{SDP}(q_2)$ but remain above $s(q_2)$ when the disappointment penalty is not large enough.

A motivation function with lower disappointment penalties is more likely to encourage the sender to use a sequential rather than a static approach to persuade. However, this does not necessarily motivate her to choose a less informative opening pitch. Specifically, with an SDP motivation function, $r_{\hat{\rho}(q_1)}(q_2)$ and $r(q_2)$ coincide. If a sequential strategy survives the SDP feature to dominate the optimal static strategy, it must also advise $p^{*b}_1 > p^{*b}_{q_1}$ because choosing

$p_1 = p^{*b}_{q_1}$ eliminates the second persuasion opportunity. In other words, for certain q_2, while a higher disappointment penalty may lower $r(q_2)$, it raises $s(q_2)$ and incentivizes the sender to adopt a more conservative opening pitch when she has determined to persuade sequentially.

Limiting or eliminating the persuasion value in the second attempt when $q_2 = p^{*b}_{q_1}$ is not the only mechanism through which a sufficiently high disappointment penalty induces a more conservative first persuasion attempt. Even if the sender can achieve the same overall persuasiveness as in Figure 4, a sufficiently large disappointment penalty can still raise the position of $s(q_2)$. This occurs because $s(q_2)$ is determined by the intersection of a fixed indifference curve $\rho(p, q_1) = \hat{\rho}$ and a family of indifference curves $\rho(p, q_2)$ indexed by the parameter q_2 where each $\rho(p, q_2)$ is tangent to the lower boundary of the feasible set. As q_2 decreases from q_1 to $p^{*b}_{q_1}$, $\rho(p, q_2)$ rotates clockwise around certain pivots. Without a disappointment penalty, as shown in Figure 4, these pivots remain within the shaded quadrilateral. However, a sufficiently high disappointment penalty can shift some of these pivots outside this region during the transition from $\rho(p, q_1)$ to $\rho(p, p^{*b}_{q_1})$, leading to a higher $s(q_2)$ for certain q_2.

This mechanism relies on the fact that a disappointment penalty imposes a greater cost on more informative strategies, reducing the effectiveness of a second persuasion attempt and pushing the sender's strategy in the first attempt away from the static optimal. This is more evident when $\mathcal{M}(p_t, q_t)$ is differentiable in p_t and when the second-order condition (S.O.C.) of the objective minimization problem is satisfied. Given these conditions, the sender's first-order condition (F.O.C.) for optimizing a sequential persuasion strategy is as follows.

$$[y(p_2^b, q_2) - q_2]\Lambda_1 + (\Lambda_2 - 1)\left(\frac{y(q_2, q_1) - q_1}{y(q_2, q_1) - q_2}\right) \leq 0, \tag{5}$$

where

$$\Lambda_2 = \frac{(q_2 - p_2^b)(y'_{q_2} + y'_{p_2^b}\frac{\partial p_2^b}{\partial q_2}) + \frac{\partial p_2^b}{\partial q_2}[y(p_2^b, q_2) - q_2]}{y(p_2^b, q_2) - p_2^b},$$

$y'_{q_2} = \frac{\partial y(p_2^b, q_2)}{\partial q_2} \leq 0$, $y'_{p_2^b} = \frac{\partial y(p_2^b, q_2)}{\partial p_2^b}$, p_2^b is a function of q_2, and Λ_1 is the first-order derivative of the objective function $\rho(p_{q_1}, q_1)$ with respect to $p_{q_1}^b$.

As $p^b = p_{q_1}^{*b}$ minimizes the objective function $\rho(p, q_1)$, Λ_1 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. It captures how the lower boundary of the feasible constraint converges to $p_{q_1}^*$ from the right. However, as an indicator of how the sender adjusts the choice of information structures to the change in the receiver's prior belief under the frictional constraint, Λ_2 is less than or equal to zero under the axiom that assumes a WDP motivation function.

When there is no disappointment penalty, such that $y'_{q_2} = 0$, Λ_2 is determined by the sender's optimal adjustment of p_2 in response to shifts in the receiver's prior belief during the second stage of persuasion. As p^b approaches $p_{q_1}^{*b}$ from the right, Λ_2 remains non-positive, capturing how $s(q_2) \in [\underline{s}(q_2), \bar{s}(q_2)]$ evolves from $p_{q_1}^*$ in Figure 4. Under Axiom 1, $y'_{q_2}(p_2^b, q_2) \leq 0$ contributes to the absolute value of Λ_2, making a higher prior belief q_2 in the second persuasion attempt not only increases the success rate of each information structure, but also reduces the disappointment penalty. Therefore, if the disappointment penalty, measured by $|y'_{q_2}(p_2^b, q_2)|$, is sufficiently large, the left-hand side of condition 5 becomes negative at $p_{q_1}^{*b}$, leading to $p_1^{*b} > p_{q_1}^{*b}$. In Figure 4, this effect is represented by raising $r(q_2)$ to the position above the lower boundary of the feasible constraint.

Proposition 4 and Corollary 2 show that SDP plays a more significant role than WDP in determining whether the sender should modify the strategy in her first persuasion attempt to secure a better subsequent opportunity to persuade. Even when the SDP feature is not necessary, sequential persuasion only becomes an intertemporal problem when the receiver exhibits sufficient impatience following a failed persuasion attempt. This requirement narrows the set of scenarios in which the sender must recognize sequential persuasion as the optimal approach. The sender needs to consider extending the persuasion game only if she perceives the receiver to be highly convinced that his belief may change after the persuasion but prone to losing patience quickly as that prospect diminishes. Otherwise, the sender's loss from treating persuasion as a static game remains relatively negligible. At the beginning of the interaction, it is crucial for the sender to assess the receiver's patience and how it evolves with the perceived likelihood of belief change. A common real-world example is when a sales representative asks about a customer's available time or willingness to listen; this serves as a crucial test for gathering information to strategically optimize the persuasion approach.

However, once there exists a $q_2 \in (0, q_1)$ such that $y(q_2, q_1) < s(q_2)$, the sender is reluctant to dramatically shift the receiver's belief in the first attempt if she anticipates a second opportunity to persuade. Adopting a more conservative approach allows her to preserve the second persuasion attempt under SDP motivation and achieves a higher subsequent persuasion value if only WDP motivation is present. In brief social interactions, making an immediate strong impression can be particularly challenging. While conventional advice often emphasizes making a good first impression, it may overlook both the difficulty in achieving this and the potential consequences of failure. Listeners may not have the time to verify an assertive introduction, and if they remain unconvinced, the introduction may come across as overconfident or even boastful. Thus, unless the sender has only one opportunity to persuade, a more conservative approach in initial interactions may be more effective. By establishing a connection rather than pressing for immediate convince, the sender reduces the risk of early rejection, keeping future persuasion opportunities open, and increasing the likelihood of a favorable impression over time.

7 Optimal Persuasion with Finite Attempts

If the sender determines that it is beneficial to extend the persuasion process and necessary to adjust the opening pitch to induce an optimal subsequent persuasion opportunity, the endogenous sequential persuasion game becomes an intertemporal optimization problem with optimal stopping. In this framework, q_t serves as the state variable, and $p_t = (q_{t+1}, y(q_{t+1}, q_t))$ becomes the control variable. The persuasion path, $(p_1, ..., p_t, ...p_\tau)$, is characterized by its length $\tau \leq T$ and the sequence of information structures along the path, both of which represent the sender's optimal decision to maximize persuasiveness. This section characterizes these properties in general settings where the attempt limit is finite. Although this extends beyond a two-stage endogenous persuasion game, its establishment is fundamentally based on the two-stage framework.

7.1 Piecemeal Information Disclosure

According to the analysis in previous sections, the signal sender will choose to extend the game duration if a persuasion strategy with an additional attempt has a higher expected persuasiveness than the original optimal design. This not only helps the sender determine whether to extend the persuasion game, but also has an important implication on the characteristic of the persuasion path $(p_1, ..., p_t, ...p_\tau)$.

Proposition 5. *Given Axiom 1 and persuasion strategies $\{\mathrm{p}^t\}_{t=1,...,\tau'}$ that are optimal for the persuasion games with $t = 1, ..., \tau'$ attempts, if p^{t+1} outperforms p^t for all $t = 1, ..., \tau' - 1$, and there exists a persuasion strategy $(p_1, ..., p_t, ...p_{\tau'+1})$ that outperforms the strategy p^τ, then in the optimal persuasion strategy with $\tau' + 1$ attempts, denoted as $\mathrm{p}^{\tau'+1} = (p_1^*, ..., p_t^*, ...p_{\tau'+1}^*)$, p_t^{*b} decreases and p_t^{*g} increases as t increases from 1 to $\tau' + 1$.*

Proof. Proof of Proposition 5. See Appendix.

A contradiction can intuitively prove Proposition 5. Given that the receiver's motivation is disappointment penalizing, in an optimal persuasion strategy with $\tau' + 1$ stages, if there are two information structures, where one of them has both the smaller p_t^{*b} and the smaller p_t^{*g}, then the sender can simply eliminate the information structure with the larger p_t^{*b} and p_t^{*g} to weakly improve persuasion effectiveness. This elimination would violate the assumption of p^{τ} outperforming all p^{t} with $t < \tau$.

This proposition introduces a "piecemeal" information disclosure strategy. The signal sender will start the persuasion by proposing a pitch with minimum information disclosure. Any signal will not cause the receiver's belief to change substantially. If previous persuasions fail, the sender will conduct another pitch that discloses more information to keep the receiver engaged and willing to be persuaded. The sender discloses more information once at a time, until the receiver is convinced or the maximum attempt is reached.

More importantly, Proposition 5 implies that as long as the sender is motivated to further extend the persuasion game for greater effectiveness, she adopts a piecemeal strategy in persuasion. Conversely, if she cannot maintain the piecemeal strategy with more attempts, she ceases to extend the persuasion path, leaving the attempt limit non-binding and path length endogenously determined by the optimal stopping. To further illustrate this, the sender's intertemporal problem of choosing information structures within the persuasion path to optimize overall persuasiveness can be formalized using a Bellman function as follows.

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

This intertemporal objective yields the transition (Euler) condition that reveals how optimal information structures evolve in a sequential persuasion strategy from stage 1 to T based on how the motivation function varies with the prior belief q_t and the targeted posterior belief p_t.

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 the information structures employed in the following periods, implying that $p^{*g}_{t+1} - p^{*b}_{t+1} > p^{*g}_t - p^{*b}_t$. Specifically, p^{*b}_t decreases and p^{*g}_t increases as t increases from 1 to T.*

Proof. Proof of Proposition 6. See Appendix.

The essential condition that underlies the piecemeal persuasion strategy in Proposition 6 coincides with the condition that motivates the sender to extend the persuasion process, as established in Proposition 2. This emphasizes the fundamental connection between an endogenously sequential persuasion strategy and its inherently piecemeal structure. Furthermore, it clarifies why these two features are closely associated. The underlying mechanism that makes a piecemeal information disclosure strategy effective is precisely the reason why the sender can improve the overall effectiveness of persuasion by adopting a sequential approach.

The receiver's inattention as a friction restricts the feasibility of implementing an ideal information structure that is characterized by small p^g_t and p^b_t. With a boundary of feasible information structures defined by $y(q_{t+1}, q_t)$ decreasing with q_{t+1}, the sender must navigate a trade-off along this boundary. She can achieve either a small p^g_t or a small p^b_t, but not both. If sequential persuasion is possible, the sender's optimal solution is to prioritize a small p^b_t, which makes a bad signal less bad. Because $y(q_{t+1}, q_t)$ also decreases with q_t, this strategy ensures that a potential failure attempt results in only a slight decrease in q_t, so the receiver's motivation to participate in subsequent persuasion attempts does not depreciate by too much. For this persuasion dynamic to remain active, the receiver's participation constraint must continue to be satisfied, which requires a progressive increase in p_t^b to accompany a progressively decreasing q_t and p_t^g along the persuasion path, inherently creating the "piecemeal" characteristic.

This active dynamic provides the sender with multiple opportunities to persuade. Because the sender only needs one successful persuasion among these many attempts, the piecemeal approach helps the sender achieve a level of persuasiveness that would be unattainable with fewer attempts on the persuasion path. In contrast, if the frictional conditions outlined in Proposition 6 are absent, the optimal static persuasion strategy dominates any feasible sequential approach. In this case, piecemeal information disclosure is not necessary and, correspondingly, the sender cannot improve persuasiveness through sequential persuasion.

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 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, ranging from 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.

7.2 The Length of Persuasion Path

When persuasion with finite attempts is formalized as a general intertemporal problem, the sender can apply the criterion defined in the two-stage problem to determine whether to extend persuasion by an additional stage, provided that the number of attempts has not reached the limit. This criterion, as established in the two-stage problem, depends on the feasible set determined by the prior belief. Therefore, the evolution of the feasible set along the persuasion path with evolving q_t determines when the sender stops persuasion under an exogenous attempt limit. Although it is challenging to explicitly characterize this correlation in general, it is possible to outline the conditions under which the optimal persuasion path length always reaches the attempt limit, serving as a benchmark. Deviations from these benchmark conditions yield scenarios with fewer optimal persuasion attempts.

Proposition 7. *Given that the initial feasible set of information structures conditional on the prior belief q_1, and the disappointment penalty is sufficiently small, then if the optimal two-stage persuasion strategy outperforms the optimal static persuasion strategy, the optimal length of the persuasion path is T, which is binding by the attempt limit.*

Proof. Proof of Proposition 7. See Appendix.

Whenever the sequential persuasion approach outperforms the static approach, a strategy involving more than one attempt can serve as a baseline. By Axiom 1, each information structure within this baseline remains feasible given the initial prior belief q_1. If the initial feasible set is convex and the disappointment penalty is sufficiently small, then any linear combination of information structures from two adjacent stages in this baseline strategy is also feasible. Under these conditions, inserting such a linear combination as an interim persuasion attempt between these adjacent stages further enhances overall persuasiveness. Thus, when the optimal persuasion strategy with t stages serves as this baseline, the sender can always improve persuasiveness by extending to $t + 1$ stages. Continuing this process, the optimal persuasion path will ultimately reach the attempt limit.

Given this, Proposition 7 reduces the determination of the optimal persuasion path length to a simpler decision: whether a two-stage approach outperforms one-shot persuasion. This extends the importance of the criterion defined in the two-stage problem to the general design of sequential persuasion and makes it central to the sender's strategy. Once the sender identifies any sequential persuasion strategy that outperforms the optimal static approach, she should exhaust all allowed persuasion attempts, adopting a finer piecemeal strategy to prevent the receiver's prior belief from evolving too quickly and causing the game to end prematurely. For the receiver, if any (exogenous) factor motivates him to stop listening to persuasion, he should take this into account and make a firm commitment to end the persuasion process at a certain point, rather than waiting for the sender to conclude.

The impact of adding an interim persuasion attempt can be visualized using an extended version of the framework depicted in Figure 2. Panel (b) of Figure 5 illustrates a scenario that permits at most three persuasion attempts ($T = 3$). In this illustration of sequential strategy, the process of determining the equivalent static persuasion strategy begins with the top endpoint, which represents the optimal strategy in the final attempt. From this

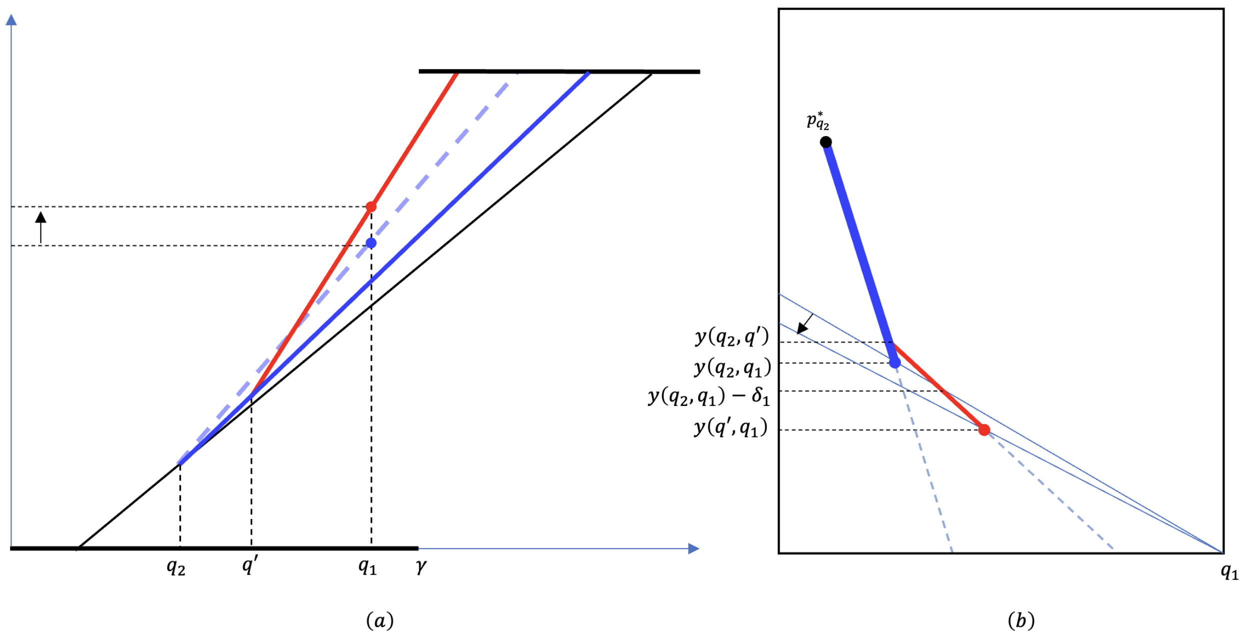


Figure 5: Value of an Additional Persuasion Attempt $(q', y(q', q_1))$

endpoint, the process moves downward along iso-value curves conditional on a certain prior belief, transitioning to another curve conditional on a larger prior belief at each connecting point until reaching the bottom endpoint on the iso-value curve conditional on the initial prior belief q_1. Conceptually, this is equivalent to pivoting the line segments from left to right around these connecting points (from q_2 to q') in panel (a), elevating the overall position of the line segment configuration (indicated by the upward shift of the red segment).

This framework can be extended to the setting with finite attempts to provide a picture of how to design a sequential persuasion strategy in general. The sender identifies a starting point on the feasible set boundary in the upper left corner of panel (b) of Figure 5, which is then connected to a sequence of line segments with decreasing slopes in order. These segments link 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.

Building on this intuition, Figure 5 echoes the insights of Propositions 5 and 6, visually showing how adding a persuasion attempt, whether by appending it to the existing path or inserting it as an interim step, improves overall persuasiveness. A steeper iso-value curve, associated with a smaller prior belief, effectively lowers the endpoint of the "wire", but its extension is limited by the horizontal coordinate set by that belief. In contrast, an iso-value curve tied to a larger prior belief can extend the "wire" further, but it is flatter and less effective in helping the "wire" reach downwards. Appending a persuasion attempt extends the "wire", while an interim attempt strikes a balance between the slope of the iso-value curve and the extent of the extension, allowing the entire "wire" to reach a lower indifference curve, which is associated with a higher persuasion value.

8 Efficiency Boundary

By extending the persuasion process, the sender can relax the frictional constraint imposed by the receiver's inattention. However, this strategy is still subject to additional constraints intrinsic to sequentiality, particularly the disappointment penalty and the attempt limit. Given an initial feasible set of information structures, the disappointment penalty further restricts sets of information structures feasible for subsequent persuasion attempts. This penalty reflects the intertemporal cost of extending the persuasion game, discouraging the sender from fully exploiting sequentiality as a means of overcoming receiver inattention. According to Axiom 1, the smallest disappointment penalty, referred to as the zero disappointment penalty, implies that the feasible set within the Bayes plausibility region does not shrink over time with previous failed persuasion attempts.

Proposition 7 indicates that the attempt limit becomes binding when the disappointment penalty is sufficiently small. Specifically, under a zero disappointment penalty, any finite attempt limit is immediately binding. In this case, provided that sequentiality is effective in mitigating the friction caused by receiver inattention, increasing the number of persuasion attempts allowed strictly enhances persuasion effectiveness. Technically speaking, with zero disappointment penalty, as the allowed number of persuasion attempts grows sufficiently large and approaches infinity, the sender's overall persuasion value approaches its efficiency limit.[3] Characterizing this efficiency boundary provides valuable insight into how much an endogenously sequential persuasion game can beat receiver inattention.

At the efficiency boundary, the constraint on the maximum number of persuasion attempts is no longer binding, even with a zero disappointment penalty. This significantly affects the sender's philosophy of designing the optimal persuasion path. When the persuasion path reaches the attempt limit, adding any new information structure requires replacing an existing one, which may not always improve overall effectiveness. In contrast, if the attempt limit is not binding, the sender is freed from this intertemporal trade-off. Any information structure can be incorporated into the path without having to substitute another. This incorporation can be beneficial even if it is not the case under a binding attempt limit. As a result, at the efficiency boundary, the sender's task in designing the optimal persuasion path reduces to determining whether a feasible information structure is qualified for inclusion, rather than selecting among qualified ones. This insight is formalized in the following lemma.

Lemma 4. *Suppose that 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 function with zero disappointment penalty, if an information structure* $(q', y(q', q_t))$ *satisfies Proposition 5, then the persuasion design with path* $(q_1, ..., q_t, q', q_{t+1}, q_{t+2}, ..., q_\tau)$ *outperforms path* $(q_1, ..., q_t, q', q_{t+2}, ..., q_\tau)$.

[3] A finite game ensures that the receiver is always willing to participate in the persuasion as long as his participation constraint is satisfied. Since the efficiency boundary is not an equilibrium, we can directly assume that the receiver will participate in the persuasion whenever his participation constraint holds. Under this behavioral assumption, the necessity for a finite attempt limit can be relaxed.

Proof. Proof of Lemma 4. See Appendix.

With zero disappointment penalty and the number of persuasion attempts allowed approaching infinity, the principle of constructing the optimal persuasion path is to incorporate as many qualified attempts as possible. Given two information structures, $(q, y(q))$ and $(q', y(q'))$ within the initial feasible set, the former is dominated by the latter if $y(q) > y(q')$ and $q > q'$. According to Lemma 4, any information structure within the initial feasible set that is not dominated by another should be included in the persuasion path.

Under this criterion, feasible information structures characterized by the smallest values of p^b or p^g, denoted by p^b_{min} and p^g_{min}, respectively, are never dominated by any other feasible information structures. Hence, they will always be part of the persuasion path that defines the efficiency boundary. According to Proposition 2, if the sequential approach improves the effectiveness of persuasion, p^b_{min} and p^g_{min} cannot appear within a single feasible information structure. Therefore, information structures featuring p^b_{min} or p^g_{min} serve as the two endpoints of the persuasion path, where $y(p^b_1, q_1) = p^g_{min}$ and $q_\tau = p^b_{min}$, $\tau \to \infty$. Based on these endpoints, the persuasion path is constructed by sequentially identifying each feasible information structure $(q_t, y(q_t, q_{t-1}))$, starting from $q_t = q_\tau$ and moving backward to $q_t = q_1$, incorporating each feasible structure whose $y(q_t, q_{t-1})$ is smaller than all those already included. This resulting persuasion path, along with the efficiency boundary it attains, is well defined and characterized, provided that the initial feasible set is convex.

Proposition 8. *With zero disappointment penalty and the attempt limit approaching infinity, when the initial feasible set of information structures is convex, all information structures* $(q_t, y(q_t, q_1))$ *with* $q_t \in [p^b_{min}, y^{-1}(p^g_{min}, q_1)]$ *are parts of the optimal persuasion path. With this* *design, the expected persuasion value is*

$$V = S + \frac{q_1 - y^{-1}(p^g_{min})}{\gamma - y^{-1}(p^g_{min})} (\gamma - S) \,,$$

where

$$S = \gamma \left[1 - \exp \left(- \int_0^{y^{-1}(p^g_{min})} \frac{dz}{y(p^b_{min} + z) - (p^b_{min} + z)} \right) \right].$$

All possible sequential persuasion designs generate expected persuasion values below this level.

Proof. Proof of Proposition 8. See Appendix.

The convexity of the initial feasible set of information structures implies that every information structure along its lower boundary, from the endpoint with p^b_{min} to the endpoint with p^g_{min}, is non-dominated and thus should be included in the optimal persuasion path. Under these conditions, the sender optimally discloses the information in the minimal increments necessary at each stage, progressively exhausting all $p^b \in [p^b_{min}, y^{-1}(p^g_{min}, q_1)]$ throughout the designed persuasion process. This persuasion path design exemplifies the finest piecemeal characteristic in the information disclosure approach.

The panel (a) of Figure 6 depicts an envelope curve formed by the collections of line segments representing individual information structures included in the optimal sequential strategy. This envelope serves as a convex "bridge", connecting two endpoints that cannot be directly linked (graphically, by a line segment) in a single information structure due to the friction of receiver inattention. In Proposition 8, S captures the additional persuasion value generated by this piecemeal disclosure design following the possible failure of the first persuasion attempt. This raises the position of this envelope curve at the position of q_1, reflecting the increased expected effectiveness of this sequential strategy compared to the static approach. This strategy approximates the persuasiveness of the information structure (p^b_{min}, p^g_{min}), although it is not feasible in a static persuasion attempt.

In a different framework, the panel (b) of Figure 6 visualizes the same "bridge", which connects p^g_{min} and p^b_{min} within the feasible set. Graphically, p^g_{min} ensures that the "bridge" extends as low as possible, while p^b_{min} provides the farthest left anchor. Additional non-dominated interim information structures 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 toward the lower left as possible, thus improving overall persuasiveness.

However, even though the sequential persuasion path can leverage both p^g_{min} and p^b_{min}, they cannot be directly associated within a single static experiment. Instead, they must be indirectly linked through intermediate persuasion stages. Graphically, the segment connecting $(p^b_{min}, y(p^b_{min}, q_\tau))$ and (q_2, p^g_{min}) is convex in panel (a) and bowed toward the origin in panel (b) of Figure 6. This curvature implies that the overall effectiveness of the sequential persuasion design, even with zero disappointment penalty, remains inferior to (p^b_{min}, p^g_{min}) as a hypothetical static information structure. Since (p^b_{min}, p^g_{min}) is weakly inferior to $(0, \gamma)$, the sequential approach is unable to recover frictionless efficiency where the receiver is fully attentive to persuasion.

8.1 Example: Efficiency Boundary When the Static Persuasion Value is $\gamma(1-\beta)$

The general efficiency boundary characterized in Proposition 8 does not specify the motivation function or the feasible set. To gain clearer insights, consider a specific example where the

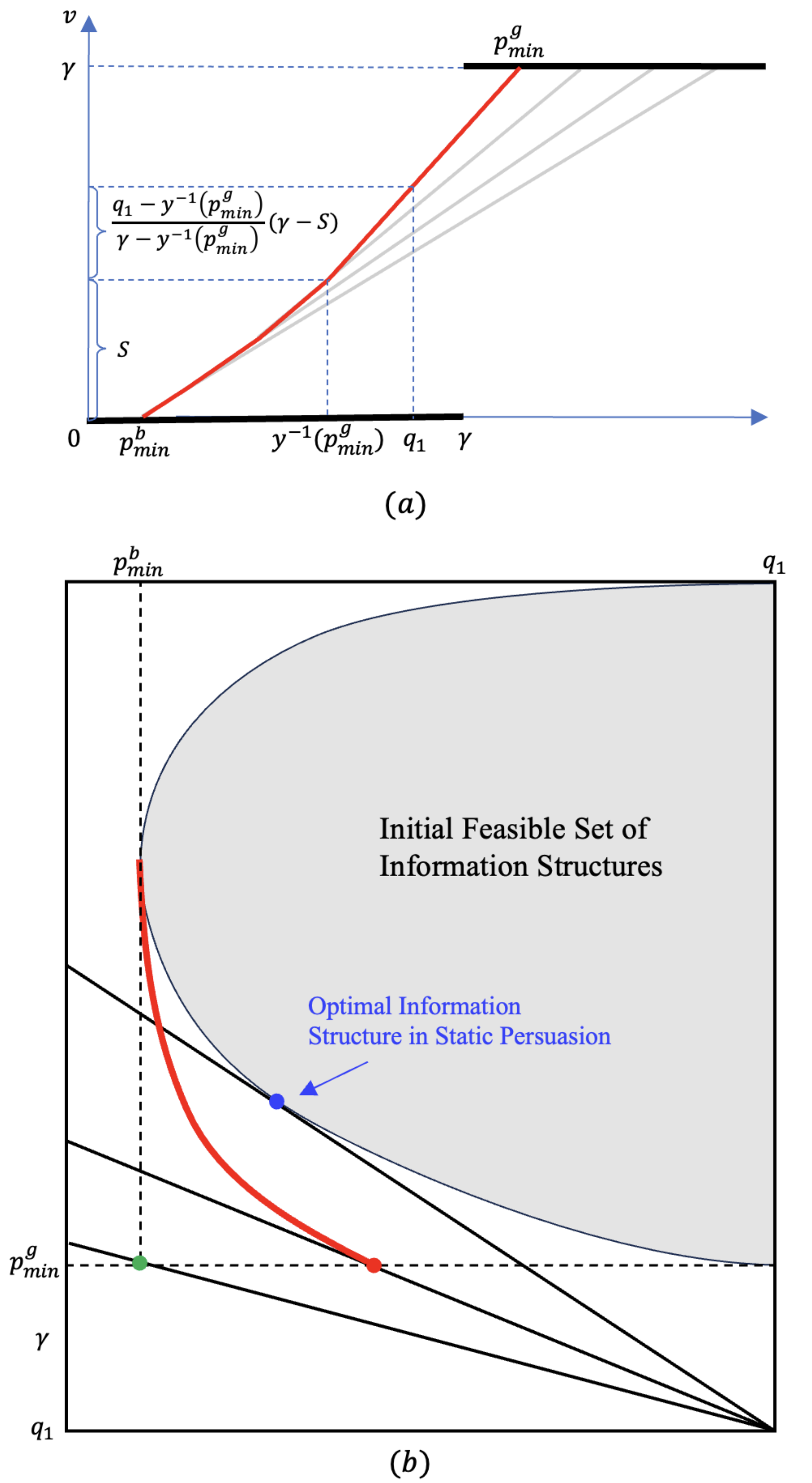


Figure 6: Efficient Boundary of Sequential Persuasion Strategies

optimal static persuasion strategy yields a persuasion value of $\gamma(1-\beta)$ in the presence of receiver inattention. In this case, the largest feasible set has its lower boundary overlap with the sender's indifference curve within the Bayes feasibility region, as shown in panel (a) of Figure 7. This feasible set defines p^b_{min}, p^g_{min}, and $y(p^b, q_1)$ explicitly in terms of exogenous parameters β, γ, and q_1, which in turn derives a clear efficiency boundary for the optimal persuasion value when the sender can strategically extend the persuasion process to mitigate the receiver's inattention as friction.

Corollary 3. *When the sender can achieve the persuasion value $\gamma(1-\beta)$ with the optimal static persuasion strategy, her optimal persuasion value with sequential approach is bounded above by $V = S + (1-\beta)(\gamma - S)$, where*

$$S = \begin{cases} \gamma \left[1 - \left(\frac{(1-\beta)(\gamma-q_1)}{q_1\beta}\right)^{1-\beta}\right] & \text{if } \beta \leq 1-q_1 \\ \gamma \left[1 - \left(\frac{\gamma-q_1}{1-q_1}\right)^{1-\beta}\right] & \text{if } \beta > 1-q_1. \end{cases}$$

Proof. Proof of Corollary 3. See Appendix.

The lower boundary of the feasible set shown in panel (a) of Figure 7 identifies $y(p^b, q_1)$ for each $p^b \in [0, q_1)$ through the relationship $\frac{y(p^b,q_1)-q_1}{y(p^b,q_1)-p^b} = \beta$. In addition, it implies that $p^g_{min} = \gamma$, $p^b_{min} = 0$ when $\beta \leq 1-q_1$, and $p^b_{min} = \frac{q_1-(1-\beta)\gamma}{\beta}$ when $\beta > 1-q_1$. Substituting these values into Proposition 8 leads directly to Corollary 3. Panel (b) of Figure 7 visualizes the efficiency boundary derived in Corollary 3 and compares it with two critical benchmarks: the frictionless efficiency level and the optimal static persuasion value.

The panel (b) of Figure 7 shows that the sequential persuasion efficiency boundary con-

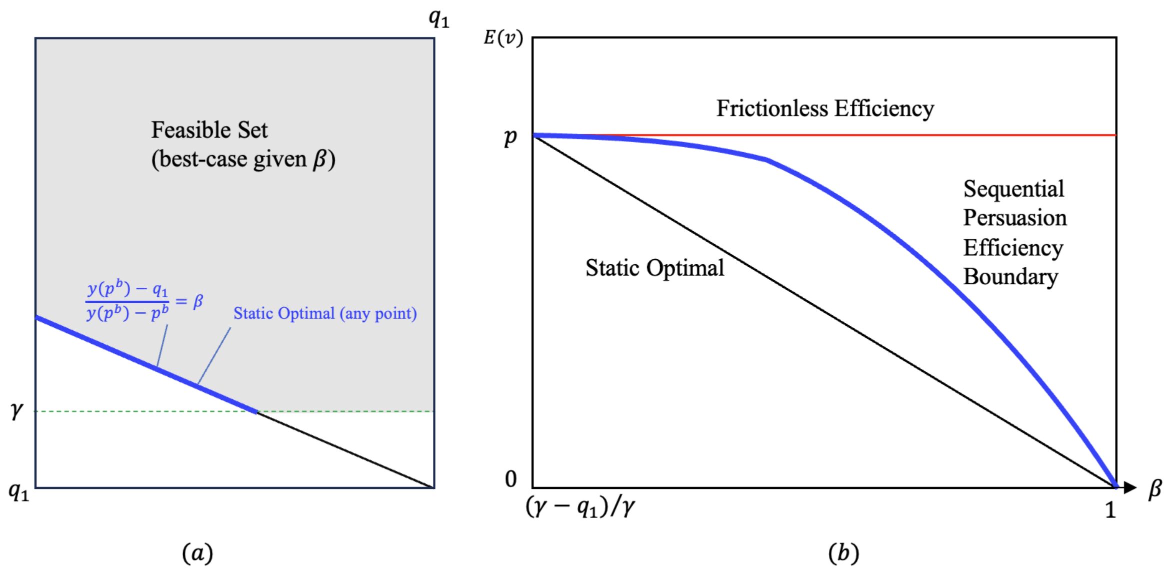


Figure 7: Efficient Boundary Conditional on the Optimal Static Persuasion Value

sistently outperforms the optimal static persuasion. Although the efficiency boundary itself is unattainable in practice, the sender's persuasion value falls within the area bounded by these two curves, with its exact position determined by deviations from the ideal conditions specified in Proposition 8 and Corollary 3. This finding underscores that sequential persuasion can effectively mitigate the friction induced by receiver inattention. However, even under ideal conditions, such as zero disappointment penalties and the receiver's infinite patience, represented by no intertemporal discounting and unlimited persuasion attempts, frictionless efficiency remains unattainable.

Moreover, the efficiency boundary evolves as β changes. β represents the chance that the sender fails in her best static persuasion attempt. Its increase reflects more severe receiver inattention and friction. The panel (b) of Figure 7 clearly illustrates that as the friction intensifies, the gap between the sequential efficiency boundary and the frictionless efficiency widens. This pattern indicates that increased friction not only impairs the effectiveness of static persuasion, but also reduces the effectiveness of sequential persuasion as a remedy.

Even if the sender can extend the persuasion process to alleviate friction in market and social interactions, efforts and policies to minimize friction directly still play an important role in improving market efficiency and social welfare.

The pattern of how the sequential persuasion efficiency boundary varies with β highlights an important threshold at $\beta = 1 - q_1$. When β crosses this threshold from below, the depreciation rate of the sequential persuasion efficiency, reflected by the absolute slope of the curve in panel (b) of Figure 7, increases sharply and overtakes the rate observed for the static optimal persuasion value. This shift occurs because when $\beta \leq 1 - q_1$, p^b_{min} stays at 0. However, for $\beta > 1 - q_1$, the sender's indifference curve is positioned above the upper left corner of panel (a) in Figure 7, causing any further increase in β to reduce p^b_{min}. This change does not occur at the tangent point that determines the static optimal strategy, so it has no effect on the static persuasion value depreciation. In contrast, the retreat of the left limit of the feasible set significantly affects the sequential approach. This distinction reveals a fundamental advantage of sequential persuasion: the sender can utilize the full boundary of the feasible set to design her strategy, rather than being confined to a single tangent point as in the static approach. Throughout this research, this mechanism underlies all key findings and fundamentally explains why sequential persuasion can achieve greater effectiveness in the presence of friction.

9 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, both of which received little attention in early and conventional persuasion research, 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 paper, 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 of receiver 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 the way 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. Furthermore, our findings suggest that the receiver's reluctance to fully engage in an initial persuasion attempt may ultimately cause him to pay more attention to subsequent persuasion if the sender adopts a sequential strategy. However, this endogenous sequential approach may not be optimal for the receiver's or social welfare. Future research could explore mechanisms that improve 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 an 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_1^g, given that q_2 is chosen as p_1^b. 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 $\tilde{p}_1^g$ 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 $(p_{\tilde{q}_2}^{*b}, p_{\tilde{q}_2}^{*g})$ 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 the sequential persuasion strategy can outperform the optimal static strategy for 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 $\rho(p_{q_1}^*, q_1) = \frac{p_{q_1}^{*g} - q_1}{p_{q_1}^{*g} - p_{q_1}^{*b}}$, 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 $q_2 = p_{q_1}^{*b}$. 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, $\{(q_2, p^g_t) | \frac{q_1 - \rho(p^*_{q_1}, q_1) q_2}{1 - \rho(p^*_{q_1}, q_1)} \leq p^g_t \leq 1, 0 \leq q_2 \leq q_1\}$. 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. To show the necessity of this proposition, assume that there exists no information structure (p^b, p^g) such that $p^b \in (0, q_1) \backslash \{p^{*b}_{q_1}\}$ and $\frac{p^{*g}_{q_1} - p^g}{p^{*b}_{q_1} - p^b} < 0$. In this case, $p^{*b}_{q_1}$ is the smallest p^b and $p^{*g}_{q_1}$ the smallest p^g among the feasible information structures when the prior belief is q_1. Otherwise, $p^*_{q_1}$ does not minimize $\rho(p, q_1) = \frac{p^g - q_1}{p^g - p^b}$. This is because according to the assumption, if there exists an information structure $p' = (p^{b'}, p^{g'})$ such that $p^{b'} < p^{*b}_{q_1}$, we must also have $p^{g'} < p^{*g}_{q_1}$, which implies $\rho(p', q_1) < \rho(p^*_{q_1}, q_1)$. The same logic applies to the situation where $p^{g'} < p^{*g}_{q_1}$.

Under Axiom 1 and given this feasible set of information structures, for any $q_2 \in (0, q_1)$, the sender can achieve $\rho(p^*_{q_2}, q_2) \geq \rho(p^*_{q_1}, q_2)$. According to $r(q_2)$ defined by Proposition 1, when $\rho(p^*_{q_2}, q_2)$ is replaced by $\rho(p^*_{q_1}, q_2)$ in Lemma 3, $r(q_2) = \psi(q_2) = p^{*g}_{q_1}$ when it is well defined. Take the derivative of $\psi(q_2)$ with respect to $\rho(p^*_{q_2}, q_2)$. Its sign is determined by $(q_2 - q_1)\rho(p^*_{q_1}, q_1) < 0$. Since $\psi(q_2)$ decreases with increasing $\rho(p^*_{q_2}, q_2)$, and $\rho(p^*_{q_2}, q_2) \geq \rho(p^*_{q_1}, q_2)$, $r(q_2) \leq p^{*g}_{q_1}$. If there is no information structure (p^b, p^g) such that $p^b \in (0, q_1) \backslash \{p^{*b}_{q_1}\}$ and $\frac{p^{*g}_{q_1} - p^g}{p^{*b}_{q_1} - p^b} < 0$, all feasible information structures with $q_2 = p^b \in (0, q_1) \backslash \{p^{*b}_{q_1}\}$ have $p^g > p^{*g}_{q_1}$, making $y(q_2, q_1) < r(q_2)$ impossible for all possible $q_2 = p^b$. Thus, we have proved the necessity of Proposition 2 by its contrapositive.

To show the sufficiency of Proposition 2 under the condition of no disappointment penalty, note that without disappointment penalty, $p^*_{q_1}$ is always feasible for any prior belief $q_2 \in (0, q_1)$. Thus, we must have $\rho(p^*_{q_2}, q_2) \leq \rho(p^*_{q_1}, q_2)$. According to the proof of necessity above, which shows that the well-defined $r(q_2)$ decreases with increasing $\rho(p^*_{q_2}, q_2)$, there must be $r(q_2) \geq p^{*g}_{q_1}$. If there exists an information structure (p^b, p^g) such that $p^b > p^{*b}_{q_1}$ and $p^g < p^{*g}_{q_1}$, $p^g < r(p^b) \leq p^{*g}_{q_1}$ is satisfied. According to Proposition 1, a sequential approach is more beneficial than the optimal static strategy. For another, if there exists a feasible information structure (p^b, p^g) such that $p^b < p^{*b}_{q_1}$ and $p^g \leq 1$, we have the same conclusion because the overall persuasion value is $1 - \rho^*_{q_1}\rho^*_{q_2}$, which is greater than $1 - \rho^*_{q_1}$ as $\rho^*_{q_2} < 1$. This finishes the proof of the sufficiency of Proposition 2 in the scenario where the disappointment penalty is sufficiently small. □

A.3 Proof of Corollary 1

Proof. Based on Lemma 3 and Proposition 1, we establish the proof of Corollary 1 formally 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, the 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, optimal two-stage approaches, conditional on different q_2, are always equivalent to the optimal static approach in Bayesian persuasion. □

A.4 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 than 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, condition $\rho(p^*_{q_2}, q_2) \geq \rho(p^*_{q_1}, q_2)$ must produce $\psi(q_2) \leq p^{*g}_{q_1}$. Hence, as long as $\{p_2 | \mathcal{M}(p_2, q_2) \geq 0\} \neq \emptyset$, it must be $r(q_2) < p^{*g}_{q_1}$.

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 outperform the optimal static persuasion strategy. To make this design possible, the chosen p_1^g must also be smaller than $\psi(q_2)$, according to Proposition 1. □

A.5 Proof of Proposition 4

Proof. When there is no disappointment penalty, a feasible information structure for a prior belief remains feasible for all possible prior beliefs.

Lemma A.1. *Given that the sender achieves the optimal persuasion value* $1-\rho_{q_1}^*$ *and* $1-\rho_{p_{q_1}^{*b}}^*$ *when the prior beliefs are* q_1 *and* $p_{q_1}^{*b}$*, respectively. For the prior belief* $q_2 \in (p_{q_1}^{*b}, q_1)$*,* $\rho(p, q_2) \geq \frac{(q_1-q_2)\rho_{p_{q_1}^{*b}}^* - (p_{q_1}^{*b}-q_2)\rho_{q_1}^*}{q_1-p_{q_1}^{*b}}$.

Proof. If Lemma A.1 is not the case, there exists a $\rho' < \rho_{p_{q_1}^{*b}}^*$ such that $\rho(p, q_2) = \frac{(q_1-q_2)\rho' - (p_{q_1}^{*b}-q_2)\rho_{q_1}^*}{q_1-p_{q_1}^{*b}}$ because $q_1 - q_2 > 0$. Then, when the sender can achieve the optimal persuasion value $1-\rho_{q_1}^*$ when the prior belief is q_1, there exists an information structure (p^b, p^g) such that $\frac{p^g-q_1}{p^g-p^b} = \rho_{q_1}^*$ and $\frac{p^g-p_{q_1}^{*b}}{p^g-p^b} = \rho'$. Since $\rho' < \rho_{p_{q_1}^{*b}}^*$, $1-\rho_{p_{q_1}^{*b}}^*$ is not the optimal persuasion value when the prior belief is $p_{q_1}^{*b}$. This contradiction validates Lemma A.1. □

Lemma A.2. *Given that the sender achieves the optimal persuasion value* $1-\rho_{q_1}^*$ *and* $1-\rho_{p_{q_1}^{*b}}^*$ *when the prior beliefs are* q_1 *and* $p_{q_1}^{*b}$*, respectively. For the prior belief* $q_2 \in (p_{q_1}^{*b}, q_1)$*,* $\rho(p, q_2) \leq \min\left\{\frac{p_{q_1}^{*g}-q_2}{p_{q_1}^{*g}-p_{q_1}^{*b}}, \frac{(1-q_2)\rho_{p_{q_1}^{*b}}^*}{1-p_{q_1}^{*b}}\right\}$.

Proof. If Lemma A.2 is not the case, for all feasible information structure, (p^b, p^g), $\frac{p^g-q_2}{p^g-p^b} > \frac{p_{q_1}^{*g}-q_2}{p_{q_1}^{*g}-p_{q_1}^{*b}}$ when $\frac{p_{q_1}^{*g}-q_2}{p_{q_1}^{*g}-p_{q_1}^{*b}} \leq \frac{(1-q_2)\rho_{p_{q_1}^{*b}}^*}{1-p_{q_1}^{*b}}$ and $\frac{p^g-q_2}{p^g-p^b} > \frac{(1-q_2)\rho_{p_{q_1}^{*b}}^*}{1-p_{q_1}^{*b}}$ when $\frac{(1-q_2)\rho_{p_{q_1}^{*b}}^*}{1-p_{q_1}^{*b}} \leq \frac{p_{q_1}^{*g}-q_2}{p_{q_1}^{*g}-p_{q_1}^{*b}}$. The former case means $(p_{q_1}^{*b}, p_{q_1}^{*g})$ is not feasible, which violates the premise that a feasible information structure for a prior belief remains feasible for all possible prior beliefs. The second case implies that feasible information structures only allow the sender to achieve the overall persuasion value of $1-\frac{(1-p^{*b}_{q_1})\rho''}{1-p^{*b}_{q_1}}$ when the prior belief is $q_2 < p^{*b}_{q_1}$, where $\rho'' > \rho^*_{p^{*b}_{q_1}}$ because $1-q_2 > 0$. Therefore, we have $\frac{p^g-p^{*b}_{q_1}}{p^g-p^b} = \frac{(1-q_2)\rho''}{1-p^{*b}_{q_1}}\frac{p^g-p^{*b}_{q_1}}{p^g-q_2}$. Take the derivative of $\frac{p^g-p^{*b}_{q_1}}{p^g-q_2}$ with respect to p^g. Its sign is determined by $p^{*b}_{q_1} - q_2 < 0$. Within the feasible set of information structures, $\frac{p^g-p^{*b}_{q_1}}{p^g-p^b}$ reaches its minimum ρ'' when $p^g = 1$. This means that the sender cannot achieve $1-\rho^*_{p^{*b}_{q_1}}$ when the prior belief is $p^{*b}_{q_1}$. These contradictions to the premises validate Lemma A.2. □

According to Lemma A.1 and Lemma A.2, for any $q_2 \in (p^{*b}_{q_1}, q_1)$, we have

$$\rho^*_{q_2} \in \left[\frac{(q_1-q_2)\rho^*_{p^{*b}_{q_1}} - (p^{*b}_{q_1}-q_2)\rho^*_{q_1}}{q_1-p^{*b}_{q_1}},\ \min\left\{\frac{p^{*g}_{q_1}-q_2}{p^{*g}_{q_1}-p^{*b}_{q_1}}, \frac{(1-q_2)\rho^*_{p^{*b}_{q_1}}}{1-p^{*b}_{q_1}}\right\}\right].$$

To obtain $s(q_2)$ such that the sequential strategy $\left((s(q_2), q_2), p^*_{q_2}\right)$ has an overall persuasiveness equivalent to $1-\hat{\rho}$ where $\hat{\rho} = \rho^*_{q_1}\rho^*_{p^{*b}_{q_1}}$, it requires $s(q_2) = \frac{\rho^*_{q_2}q_1-\hat{\rho}q_2}{\rho^*_{q_2}-\hat{\rho}}$. According to this expression, the derivative of $s(q_2)$ with respect to $\rho^*_{q_2}$ has the sign determined by $(q_2-q_1)\hat{\rho}$, which is negative because $q_2 < q_1$. Therefore, the upper bound of $s(q_2)$, $\bar{s}(q_2)$, is $\frac{\rho^*_{q_2}q_1-\hat{\rho}q_2}{\rho^*_{q_2}-\hat{\rho}}$ when $\rho^*_{q_2} = \frac{(q_1-q_2)\rho^*_{p^{*b}_{q_1}} - (p^{*b}_{q_1}-q_2)\rho^*_{q_1}}{q_1-p^{*b}_{q_1}}$. The lower bound of $s(q_2)$, $\underline{s}(q_2)$, is $\frac{\rho^*_{q_2}q_1-\hat{\rho}q_2}{\rho^*_{q_2}-\hat{\rho}}$ when $\rho^*_{q_2} = \min\left\{\frac{p^{*g}_{q_1}-q_2}{p^{*g}_{q_1}-p^{*b}_{q_1}}, \frac{(1-q_2)\rho^*_{p^{*b}_{q_1}}}{1-p^{*b}_{q_1}}\right\}$. The definitions of $\bar{s}(q_2)$ and $\underline{s}(q_2)$ here prove Proposition 4. □

A.6 Proofs and Discussion of Propositions 5, 6, and 7

Proof. According to the model setting, the receiver will pay attention to the persuasion only when at least one of the signals causes his belief in the state $\omega = h$ to exceed γ. In this case, the persuasion continues only when the signal reinforces the receiver's current choice of "not act". Therefore, in a persuasion path reflecting the sender's optimal design, $q_{t+1} = p^b_t$ decreases as t increases.

Based on this fact, if for any τ' we have $p^g_{\tau'+1} < p^g_{\tau'}$, then $p^\theta_{\tau'+1} < p^\theta_t$ for both $\theta = g$ and $\theta = b$. Given that Axiom 1 holds, eliminating the information structure $p_{\tau'}$ will increase the relevant prior belief from q_t to $q_{t-1} > q_t$ and will not change the feasibility of $p_{\tau'+1}$ and improve the overall persuasiveness. However, this contradicts the premise of this proposition that the persuasion path with length $\tau' + 1$ outperforms all the persuasion paths with length $1, 2, ..., \tau'$. This contradiction proves Proposition 5. □

Proposition 5 and Proposition 6 have the same claim with different conditions. In Proposition 6, the condition of optimal persuasion with a certain path outperforming any persuasion design with a shorter path is replaced by the shape of the feasible set of information structures.

Proof. When the sender's intertemporal objective is formalized using Bellman equation:

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

Taking the derivative of this equation with respect to q_{t+1} yields the F.O.C, or Euler condition:

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

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, and $p^{*\theta}_t$, $\theta = b, g$, represents the optimal information structure at the stage t in this persuasion path design. Grouping terms according to $y'_{t,q_{t+1}}$ and $y'_{t+1,q_{t+1}} - 1$, and leveraging the fact that $q_{t+1} = p^{*b}_t$ yields:

$$\frac{1}{p^{*g}_{t+1} - p^{*b}_{t+1}} = \frac{1}{p^{*g}_t - p^{*b}_t} + y'_{t,q_{t+1}} \frac{q_t - q_{t+1}}{(p^{*g}_t - q_t)(p^{*g}_t - q_{t+1})} + [y'_{t+1,q_{t+1}} - 1] \frac{q_{t+1} - q_{t+2}}{(p^{*g}_{t+1} - q_{t+1})(p^{*g}_{t+1} - q_{t+2})}, \tag{A.2}$$

When $y'_{t,q_{t+1}} < 0$ and $y'_{t+1,q_{t+1}} < 1$, we have $\frac{1}{p^{*g}_{t+1} - p^{*b}_{t+1}} < \frac{1}{p^{*g}_t - p^{*b}_t}$ according to (A.2), given that $q_t > q_{t+1} > q_{t+2}$. This implies $p^{*g}_{t+1} - p^{*b}_{t+1} > p^{*g}_t - p^{*b}_t$, 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:

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

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

Proposition 7 is closely related to Propositions 5 and 6, which implies that on an optimal persuasion path, $q_{t+1} = p^b_t$ decreases and p^g_t increases as t increases.

Proof. Since $q_{t+1} = p^b_t$ decreases as t increases, q_1 is greater than any other prior belief on the persuasion path. According to Axiom 1, all information structures on the persuasion path remain within the initial feasible set of information structures where the prior belief is q_1. Given that this feasible set is convex, for any adjacent information structures on the path, p_{t+1}, p_t, their linear combination, p' is also within the initial feasible set.

When the disappointment penalty is sufficiently small, p' remains feasible conditional on the prior belief $q_t = p^b_{t-1}$. Therefore, for any optimal persuasion path with a length greater than 1, it is always possible to find such an information structure p'. When including this information structure between the original stages t and $t + 1$, the overall effectiveness of persuasion improves with a sufficiently small disappointment penalty that does not affect the feasibility of p_{t+1}. To see this, $\frac{p^g_{t+1} - p'^b}{p^g_{t+1} - p^b_{t+1}} \frac{p'^g - p^b_t}{p'^g - p'^b} < \frac{p^g_{t+1} - p^b_t}{p^g_{t+1} - p^b_{t+1}}$ always holds if $p^b_{t+1} < p'^b < p^b_t$ and $p^g_{t+1} > p'^g > p^g_t$ are satisfied.

The above result implies that the optimal persuasion design conditional on a certain length outperforms the optimal designs for the shorter path that is longer than 1 stage. Therefore, under the conditions outlined in Proposition 7, if any sequential persuasion strategy outperforms the static optimal persuasion strategy, the sender will exhaust all persuasion attempts allowed in her design of the persuasion strategy, which completes the proof of Proposition 7. □

According to these proofs, Propositions 5, 6, and 7 together provide a systematic approach to predicting the optimal sequential persuasion design with more than two stages, while still grounding the analysis in the comparison between the optimal static and two-stage persuasion strategies. If the optimal two-stage persuasion strategy outperforms the static approach, it establishes a baseline that includes at least two information structures, allowing Proposition 7 to function, guiding the sender to further extend the persuasion process. The argument in Proposition 7 then serves as the premise for Proposition 5, which delineates the piecemeal structure of the sequential persuasion design, allowing Proposition 7 to reengage. In essence, the analysis of the two-stage scenario serves as a catalyst for a chain reaction, ultimately driving the sequential approach to exhaust the available attempt limit.

A.7 Proofs of Proposition 8 and Corollary 3

Proof. Assume a zero disappointment penalty and suppose that the sender has infinite chances to persuade the receiver, and in each chance, the receiver will pay attention as long as his participation constraint is satisfied. According to Lemma 4, the persuasion path that represents the efficiency boundary is $\{(p_t^b, p_t^g)\}_{t\in[1,2,\ldots]}$, where all $p^b \in [p^b_{min}, y^{-1}(p^g_{min}, q_1)]$ are within this path, and $p_t^g = y(p_t^b, q_1)$. In this persuasion path, the sender chooses the information structure $(p_1^b, p_1^g) = (y^{-1}(p^g_{min}, q_1), p^g_{min})$. Let S be the persuasion value following this stage conditional on that the persuasion fails on this stage. The overall persuasion value of this persuasion path is $V = S + \frac{q_1 - y^{-1}(p^g_{min})}{\gamma - y^{-1}(p^g_{min})}(\gamma - S)$.

Consider the approach of "backward induction", in the last stage of the persuasion path, the sender has the prior belief $p^b_{min}+\delta$ and chooses the information structure $(p^b_{min}, y(p^b_{min}, q_1))$. Accordingly, the sender's persuasion value in this last stage is $v_0 = \delta\frac{\gamma}{y(p^b_{min},q_1)-p^b_{min}}$. Without loss of generality, suppose that the sender always follows the same rule that whenever the prior belief is q, she chooses the information structure $(q-\delta, y(q-\delta, q_1))$. Based on the persuasion value v_0 and any v_ι, $\iota = 0, 1, 2, ...$, $v_{\iota+1}$ is represented by $v_{\iota+1} = \delta\frac{\gamma+\sum_{k=0}^{\iota} v_k}{y(p^b_{min}+\iota\delta,q_1)-(p^b_{min}+\iota\delta)}$. In order to ensure that the sender exhausts all $p^b \in [p^b_{min}, y^{-1}(p^g_{min}, q_1)]$ along the path to maximize the persuasion value, δ should approach 0. Therefore, the value of S is given by $\lim_{\delta\to 0}\sum_{\iota=0}^{\frac{y^{-1}(p^g_{min},q_1)}{\delta}} v_\iota$.

To obtain the closed-form value of S, consider $S_{\iota+1} = S_\iota + v_{\iota+1}$. Substituting into the expression of $v_{\iota+1}$, we have

$$\frac{S_{\iota+1} - S_\iota}{\delta} = \frac{\gamma + S_\iota}{y(p^b_{min} + (\iota+1)\delta, q_1) - (p^b_{min} + (\iota+1)\delta)}. \tag{A.3}$$

Consider that S is a function of ι, then when $\delta \to 0$, the LHS of (A.3) approximates the derivative of S with respect to ι. This condition (A.3) can be understood as an ordinary differential equation:

$$S'(z) = \frac{\gamma + S(z)}{y(p^b_{min} + z, q_1) - (p^b_{min} + z)},\ S(0) = 0$$

Note that $S'(z) = \frac{d(\gamma+S(z))}{dz}$. Solving this ordinary differential equation for $S(z)$ yields:

$$S = \gamma\left[1 - \exp\left(-\int_0^{y^{-1}(p^g_{min})}\frac{dz}{y(p^b_{min}+z) - (p^b_{min}+z)}\right)\right]. \tag{A.4}$$

This finishes the proof of Proposition 8. □

Based on Proposition 8, substituting the functional form of $y(\cdot, q_1)$ and the values of p^b_{min} indicated by the lower boundary of the feasible set that overlaps the indifference curve associated with the static persuasion value $\gamma(1-\beta)$ produces the proof of Corollary 3.

Proof. As indicated in panel (a) of Figure 7, p^g_{min} is fixed at γ for all levels of persuasion value. Based on this fact, when $\beta \leq 1-q_1$, we have $p^b_{min} = 0$, $y(z) = \frac{q_1-\beta z}{1-\beta}$, and $y^{-1}(\gamma, q_1) = \frac{q_1-(1-\beta)\gamma}{\beta}$; when $\beta > 1-q_1$, $p^b_{min} = \frac{q_1-(1-\beta)}{\beta}$, $y(p^b_{min}+z) = \frac{q_1-(p^b_{min}+z)\beta}{1-\beta}$, and $y^{-1}(\gamma, q_1) = \frac{q_1-(1-\beta)\gamma}{\beta}$. Substituting these values into A.4 produces:

$$S = \begin{cases} \gamma\left[1-\left(\frac{(1-\beta)(\gamma-q_1)}{q_1\beta}\right)^{1-\beta}\right] & \text{if } \beta \leq 1-q_1 \\ \gamma\left[1-\left(\frac{\gamma-q_1}{1-q_1}\right)^{1-\beta}\right] & \text{if } \beta > 1-q_1. \end{cases}$$

This finishes the proof of Corollary 3. □