University of Toronto
Department of Economics

Working Paper 791

Persuading while Learning

By Itai Arieli, Yakov Babichenko, Dima Shaiderman and Xianwen Shi

January 16, 2025

Persuading while Learning

Itai Arieli,* Yakov Babichenko,† Dimitry Shaiderman,‡ Xianwen Shi§

January 2, 2025

Abstract

We propose a dynamic persuasion model of product adoption, where an impatient, long-lived sender commits to a dynamic disclosure policy to persuade a sequence of short-lived receivers to adopt a new product. The sender privately observes a sequence of signals, one per period, about the product quality, and therefore the sequence of her posteriors forms a discrete-time martingale. The disclosure policy specifies ex ante how the sender's information will be revealed to the receivers in each period. We introduce a new concept called "Blackwell-preserving kernels" and show that if the sender's belief martingale possesses these kernels, the family of optimal strategies for the sender takes an interval form; namely, in every period, the set of martingale realizations in which adoption occurs is an interval. Utilizing this, we prove that if the sender is sufficiently impatient, then under a random walk martingale, the optimal policy is fully transparent up to the moment of adoption; namely, the sender reveals all the information she privately holds in every period.

1 Introduction

In many practical applications of information design, the informed party (the sender) is not always fully aware of the realized state of the world. This information may be gradually revealed to her over time. For example, consider a government during the COVID-19 pandemic who wants to maximize early vaccination rates despite uncertainties about whether the vaccine is *moderately* effective or *highly*

*Faculty of Data and Decision Sciences, Technion, E-mail: iarieli@technion.ac.il.
†Faculty of Data and Decision Sciences, Technion, E-mail: yakovbab@technion.ac.il.
‡Faculty of Data and Decision Sciences, Technion, E-mail: dima.shaiderman@gmail.com.
§Department of Economics, University of Toronto, E-mail: xianwen.shi@utoronto.ca.

effective. The government prefers widespread vaccination for the collective benefit, even if the vaccine is only moderately effective. However, individuals are willing to be vaccinated only if their belief that the vaccine is highly effective exceeds a certain threshold. The government periodically receives updates about vaccine effectiveness from ongoing research and clinical trials, and must decide how to share this sequentially arriving information with the public.

Similar scenarios include sellers introducing new products to potential buyers, central banks releasing stress test results to prevent bank runs, or central banks sharing foreign exchange market information to defend a currency peg. In all these applications, the sender must decide how to reveal information to the receiver while simultaneously learning about the underlying state, and the interests of the sender and the receiver are not aligned.

If the sender can ex ante commit to a dynamic disclosure policy, how should the sender reveal the information she privately learns over time? What is the central tradeoff faced by the sender in deciding whether to reveal more information today or withhold it until tomorrow? When is full transparency in every period optimal for the sender?

To address these questions, we frame our dynamic persuasion model within the context of product adoption. In this model, a long-lived sender commits in advance to a dynamic disclosure policy to persuade a sequence of short-lived receivers to adopt her new product. The quality of the product can be either high or low, and is initially unknown to all parties. Each period, the sender receives a signal about the product quality and follows the committed disclosure policy to partially reveal what she has learned so far. In each period, a receiver (or a cohort of receivers) arrives, observes both current and past revelations, and decides whether to adopt the product. The sender aims to maximize early adoption rates regardless of the product quality, while the receivers will adopt the product only if their belief that the quality is high exceeds a certain threshold.[1]

The sender in our model faces a dilemma: prioritizing today's adoption rate may compromise tomorrow's success rate. One approach is the *greedy policy*, where the seller chooses information disclosure in each period to maximize the probability of adoption for that period. Conversely, the sender can wait, revealing no information initially, and then leverage the accumulated information later. As

[1]Equivalently, our model can be viewed as an interaction between a long-lived sender and a single long-lived receiver; once the receiver's posterior belief has exceeded the desired threshold for adoption, the sender stops revealing information and the long-lived receiver adopts immediately.

we will demonstrate, both policies can be optimal. More broadly, a significant aspect of our analysis is the interplay between revealing information to persuade today or waiting to learn more about the state and persuade tomorrow.

To gain insight for our analysis, it is useful to view our sender's problem as an *optimal stopping problem*. Note that the sender's learning process can be fully captured by a sequence of her posteriors or a belief martingale. Given that each receiver's action choice is binary, we can restrict the sender to issuing binary recommendations: "adopt" or "do not adopt." A disclosure policy specifies, for each period and for every privately observed sender's posterior, the probability of sending an "adopt" recommendation. Since all current and past recommendations are public, once a receiver follows an "adopt" recommendation, the sender will stop releasing further information by issuing only "adopt" recommendations in all future periods. All subsequent receivers have the same information as the receiver who first adopts and hence will adopt as well. Therefore, if the sender is restricted to use obedient disclosure policies where her recommendations are always followed, then the game essentially stops as soon as the first "adopt" recommendation is sent out.

Even though the sender's problem can be viewed as a stopping problem, it is more complex than the standard stopping problem with finite-dimensional state variables. The complication arises because the state variables here are positive measures which are generally infinite-dimensional. Without further restrictions to the sender's belief martingale, it is well known in the literature that such problems become intractable as soon as the number of states exceeds three (see the literature review below for a related discussion).

To provide a more tractable framework, we reformulate the sender's optimization problem so that the state variables are positive measures. This identification is achieved by associating the state at any point in time with the remaining total probability mass available to the sender for persuading the receiver. Our first main contribution is the introduction of a new concept called *Blackwell-preserving kernels*. This concept imposes a natural restriction on the sender's belief martingale, which is crucial to our analysis and may prove useful in other applications involving martingales.

Recall that a probability kernel specifies the belief distribution for the next period based on the sender's current belief, defining a mapping between probability measures. We call a probability kernel Blackwell order-preserving if it maintains the Blackwell order on probability measures. Simply put, if one sender's belief distribution today is Blackwell more informative than another sender's and the belief transition is governed by this kernel, the more informed sender will retain this advantage tomorrow. Any belief martingale can be generated using a sequence of probability kernels. A martingale is Blackwell-order preserving if each of its probability kernels is Blackwell-order preserving. We demonstrate that several standard martingales, such as those generated by conditionally independent signals or those inducing a random walk, are Blackwell-order preserving.

Our second main contribution is to show that if the sender's belief martingale has Blackwell-preserving kernels, the class of optimal strategies for the sender has a simple structure called *interval policies* (Theorem 1). Interval policies are a class of policies in which the sender issues "adopt" recommendation if and only if the posterior belief lies in a time-dependent closed interval around the desired belief threshold for adoption.[2] With this characterization, the sender only eeds to choose the size of an interval around the adoption threshold for each period, which essentially reduces the sender's infinite-dimensional problem to a one-dimensional one, making the problem much more tractable.

Our third contribution demonstrates how the characterization of interval policies can be used to derive optimal dynamic disclosure policies in various learning environments. In particular, to approximate the motivating COVID-19 application, we consider the case where the sender's belief martingale is formed by a random walk on the grid and the sender's payoff is a discounted sum of her instantaneous gains from adoption. We prove that as long as the sender's discount factor lies below a grid-dependent threshold, it is optimal for the sender to act greedily in each period (Theorem 2). This provides a partial characterization of conditions under which full transparency in every period is optimal for the sender.

1.1 Related Literature

This paper contributes to the growing literature on dynamic persuasion which extends the classical model of Bayesian persuasion ([Aumann et al., 1995, Kamenica and Gentzkow, 2011]) to various dynamic settings.

One branch of literature focuses on a setting in which the state of the world evolves as a Markov chain, and the sender privately observes the state in each pe-

[2] The optimality of interval policies has also been established in [Guo and Shmaya, 2019, Guo et al., 2022b]. However, their models are very different from ours: their receiver is privately informed, their sender discloses information only once, and they need the likelihood ratio dominance for the optimality of the interval policies.

riod. [Renault et al., 2017] characterize cases in which the greedy policy (namely, persuading optimally in each period) is optimal for a sender with discounted utilities. This includes Markov chains with binary states and renewing Markov chains. Our model is also a Markovian model once we identify the Markov chain's states with the posteriors of the sender. [Renault et al., 2017] demonstrate that the problem becomes involved already for ternary states, while the above reduction creates a Markov chain with as many states as the number of possible posteriors of the sender along the learning process. This number is large (or even infinite) in most of our applications and for this reason, their result is not directly applicable in our setting. [Lehrer and Shaiderman, 2021] study, in the same setting, when is the statically optimal value achievable in the dynamic setting. A continuous-time analog of this setting, including some important binary-state examples has been studied by [Ely, 2017].

Several papers such as [Henry and Ottaviani, 2019, Orlov et al., 2020, Bizzotto et al., 2021, McClellan, 2022] study models in which, similarly to ours, the sender is initially uninformed and dynamically learns the state. These models assume that the information acquired by the sender is *public* to the sender and the receiver. In contrast in our model, the sender receives costless information regardless of the information he reveals to the receiver. Thus in contrast to these models, we allow the sender to "accumulate" information.

A crucial aspect of our model is that the dynamically learned information is exogenous and *private* to the sender. In particular, we assume that the learning procedure is independent of the experiment that the information is revealed to the receiver.[3]

Different sender's objectives such as incentivizing exploration [Kremer et al., 2014, Che and Hörner, 2018], maximizing suspense and surprise [Ely et al., 2015], or maximizing effort [Ely and Szydlowski, 2020] have been studied in a dynamic persuasion setting in the case where the sender is initially fully informed about the state.

Several papers have developed models of dynamic persuasion to provide jus-

[3]We note that in the COVID example, one may argue that the number of vaccinated people also affects the information available to the state. While this is true, we neglect this effect in our analysis. Thus, we assume that information about the vaccine will be released (say, from other sources) regardless of the number of vaccinated individuals. See [Guo et al., 2022a] for a two-period model in which a seller uses dynamic pricing to affect a buyer's learning based on their first-period consumption. The feedback the buyer receives increases in the amount consumed in the first period.

tifications for the sender-optimal equilibrium outcome with commitment in the model of static persuasion. For example, [Che et al., 2023] show in a model of dynamic persuasion where information generation and processing are both costly for the sender and the receiver that the sender-optimal outcomes in [Kamenica and Gentzkow, 2011] can be approximated in a Markov perfect equilibrium as the persuasion cost vanishes.[4] [Best and Quigley, 2024] and [Mathevet et al., 2022] show that, in a model where a long-lived sender plays a cheap-talk game against a sequence of short-lived receivers, the sender-optimal equilibrium can be supported by reputation.

Finally, besides the persuasion literature, there is also the literature on the disclosure of verifiable information in dynamic settings. This includes [Au, 2015, Knoepfle and Salmi, 2024]. Notice that the disclosure problem restricts the sender's strategies to the timing of revealing any evidence while the persuasion problem allows for much richer policies for the sender which include a partial revelation of these pieces of information and their timing.

2 The Model

In the section, we first set up the model. Then we provide two useful perspectives of studying the model. In the first perspective, we view the persuasion problem as an optimal stopping problem. This perspective is very useful conceptually, but to solve our problem, we need the second perspective which is based on dynamic programming with measures and sub-measures as state and control variables.

2.1 Model Setup

Consider a dynamic persuasion model where a long-lived sender (she) interacts with a sequence of short-lived receivers (each, he) over T periods. The number of periods can be either finite ($T \in \mathbb{N}$) or infinite ($T = \infty$). There is an unknown state $\omega \in \{0, 1\}$ which is randomly drawn according to a common prior $\pi \in (0, 1)$ with $\pi = \mathbb{P}(\omega = 1)$.

In each period $t = 1, 2, ..., T$, the sender privately observes a noisy signal about ω and chooses what to disclose to the receiver arriving in that period. The receiver at time t must choose an action a_t from a binary action set $A = \{0, 1\}$, where

[4] See [Honryo, 2018] for an earlier contribution to dynamic persuasion with persuasion cost, and [Escudé and Sinander, 2023] for a model of slow persuasion where the sender is restricted to a graduality constraint.

action $a_t = 1$ is called *adoption*. The utility of the receiver at time t depends on his action a_t and the state ω:

$$u(\omega, a_t) = u_t(\omega, a_t) = \begin{cases} 0 & \text{if } a_t = 0 \\ 1 & \text{if } a_t = \omega = 1 \\ -\frac{l}{1-l} & \text{if } a_t = 1, \omega = 0. \end{cases}$$

Therefore, the receiver at time t prefers adoption if and only if his belief (of $\omega = 1$) lies in $[l, 1]$. To rule out the trivial case, we assume $l > \pi$.

Upon receiving each noisy signal, the sender updates her posterior belief about the state. We will directly focus on the sequence of her posteriors rather than the sequence of noisy signals that she receives. Let $x_t \in [0, 1]$ denote the probability that the sender assigns to state $\omega = 1$ at time t, and identify the set of the sender's beliefs with the interval $[0, 1]$. Specifically, we assume that the process according to which the sender learns about the state is governed by a (commonly known) martingale $\mathbf{X} = (X_t)_{t=0,1,...,T}$ supported on the interval $[0, 1]$ with X_0 being the Dirac measure on the common prior π.[5] We restrict attention to cases in which $\mathbf{X}$ is a *Markovian* martingale. That is, the behavior of the martingale from time $t + 1$ on depends only on the realization of X_t (rather than the realization of the entire history $(X_1, ..., X_t)$).

The sender prefers the receivers to adopt regardless of the state realization. We assume that the sender can ex ante commit to an information revelation policy which specifies, in each period t, for each realization of the sender's posteriors, whether and how to reveal to the receiver what the sender has learned so far. In each period $t = 1, ..., T$, the receiver observes both past and current revelations and forms his posterior belief. The belief sequence of the receivers also forms a belief martingale. The martingale property implies that, if the sender stops revealing information from period t onwards, the receiver's belief will remain constant in all future periods. Therefore, if the sender successfully persuades the receiver at time t to adopt, she can ensure adoption in all subsequent periods by withholding information from period t onwards.

To complete the model description, we formally specify the sender's payoff from the receivers' actions. The observation we have made in the previous paragraph implies that, in defining the sender's payoff, the only relevant sequences of actions take the following form: $(a_t)_{t=1,2,...,T} = (0, ..., 0, 1, 1, 1, ...) \in \{0, 1\}^T$. Let w_t be

[5] We follow the convention that bold upper case letters denote vectors of random variables, upper case letters denote random variables, and low case letters denote their realizations.

the sender's "accrued" future payoffs when the first adoption occurs at time t. We assume that the sequence $\{w_t\}$ is decreasing in t, capturing the idea that the sender prefers an early first adoption to a late adoption. For example, we can take $w_t = \delta^{t-1}$ for the case where the sender's payoff is the δ-discounted utilities. Alternatively, we may interpret $w_t - w_{t+1}$ as the mass of receivers who arrive at time t (with $w_{T+1} = 0$), and thus the sender's payoff equals the expectation of $\sum_{t=1}^{T}(w_t - w_{t+1})a_t$.

Remark 1. The current model assumes that receivers are short-lived and they arrive sequentially. Mathematically, it is isomorphic to an alternative model where there is a single receiver who is long-lived but myopic. In fact, since the sender has full commitment, our analysis remains valid if there is a single receiver who is long-lived and strategic. The sender can commit to no further information revelation as soon as the receiver's posterior belief exceeds the adoption cutoff, and hence once the sender stops revelation, the receiver has incentive to adopt in the current and all future periods. □

2.2 Persuasion as a stopping problem

An information revelation policy is a sequence of mappings from realizations of the martingale $(x_1, ..., x_t)$ to distributions over abstract signals, $t = 1, ..., T$. Given that the action choice of each receiver is binary, the standard direct revelation argument implies that we can replace the abstract signals by a binary recommendation: "adopt" or "do not adopt". Moreover, as we have argued earlier, once the first "adopt" recommendation is sent out and followed at time t, the sender will reveal no further information by sending only the "adopt" recommendation in the subsequent periods, and all future receivers will adopt. Therefore, an information revelation policy is simply a stopping time τ on the martingale $\mathbf{X}$, specifying when to send out the first "adopt" recommendation and essentially stop the game.

Formally, we identify the sender's policy with a *randomized* stopping rule defined in terms of a random time τ and measurable mappings $\tau_t : [0,1]^t \to [0,1]$, $t = 1, \ldots, T$, so that

$$\mathbb{P}(\tau = t | \tau \geq t, X_1 = x_1, \ldots, X_t = x_t) = \tau_t(x_1, \ldots, x_t) \tag{1}$$

That is, conditional on no "adopt" recommendation being sent prior to t, the probability of sending the first "adopt" recommendation (i.e., stopping) is determined by the τ_t function. Moreover, we require our stopping rule to satisfy an *obedience constraint* in the form of:

$$\mathbb{E}[X_t|\tau = t] \geq l, \quad \forall t = 1, ..., T. \tag{2}$$

By Bayes rule, upon receiving an "adopt" recommendation at time t, the receiver's posterior belief is $\mathbb{P}(\omega = 1|\tau = t) = \mathbb{E}[X_t|\tau = t]$. Therefore, constraint (2) requires that the receiver is (weakly) better off by following an "adopt" recommendation. Denote by $\mathcal{T}$ the set of all randomized stopping rules τ satisfying the obedience constraint (2).

The obedience constraint (2) must bind in any optimal solution because, otherwise, the sender can strictly increase the probability τ_t of sending the first "adopt" recommendation while still meeting the obedience constraint. Our assumption of $\pi < l$ implies that $\mathbb{E}[X_t|\tau > t] < l$, and thus the first adoption occurs before time t if and only if $\tau \leq t$. The sender's optimization problem is thus[6]

$$V^* = \sup_{\tau \in \mathcal{T}} \left\{ \sum_{t=1}^{T} \mathbb{P}[\tau = t] w_t \right\}. \tag{3}$$

Example 1 (Two Periods, Binary Signals)**.** To demonstrate the subtleties in the above persuasion problem, consider a two-period model where the sender is sequentially exposed to two conditionally independent signals S_1 and S_2, each with binary support $\{L, H\}$, with one signal in each period. The state is equally likely ex ante, i.e., $\pi = 1/2$. For $t = 1, 2$, the distribution of signal S_t conditional on state ω is $\mathbb{P}(S_t = H|\omega = 1) = \mathbb{P}(S_t = L|\omega = 0) = q_t$ with $q_t \in [1/2, 1]$. Let

$$x_s = \mathbb{P}(\omega = 1|S_1 = s), s \in \{L, H\}$$

be the sender's period-one posterior estimate of the state after receiving signal $S_1 = s$ and

$$x_{ss'} = \mathbb{P}(\omega = 1|S_1 = s, S_2 = s'), s, s' \in \{L, H\}$$

be the period-two posterior estimate after receiving signals $S_1 = s$ and $S_2 = s'$. Then the martingale $\mathbf{X} = (X_0, X_1, X_2)$ is given by $\text{supp}(X_0) = \{\pi\}, \text{supp}(X_1) = \{x_L, x_H\}$, and $\text{supp}(X_2) = \{x_{LL}, x_{LH}, x_{HL}, x_{HH}\}$. See Figure 1 for illustration. We normalize the sender's utility by setting $w_1 = 1$.

Since there are only two periods, the optimal policy must be greedy at $t = 2$. That is, in the second period the sender persuades the receiver with the maximal possible probability.

[6]For the case $T < \infty$, the supremum can be replaced by a maximum using backward-induction arguments. For the case $T = \infty$, we prove the existence of the maximum for all the specific applications that we consider.

Is greedy policy also optimal at $t = 1$? Suppose that $l = 18/25$, the information at the first period is generated by a signal with precision $q_1 = 3/4$ and in the second period by a conditionally independent signal with precision $q_2 = 4/5$. Then for $w_2 \in [0, 0.618]$,[7] the optimal policy is greedy with an adoption/stopping probability of $25/47$ in $t = 1$. If $w_2 \in [0.955, 1]$, the optimal policy stays mute (i.e., with zero probability of adoption) at time $t = 1$. If $w_2 \in (0.618, 0.955)$, the optimal policy induces adoption with probability $25/58$ at time $t = 1$, which is smaller than the adoption probability of $25/47$ under the greedy policy.

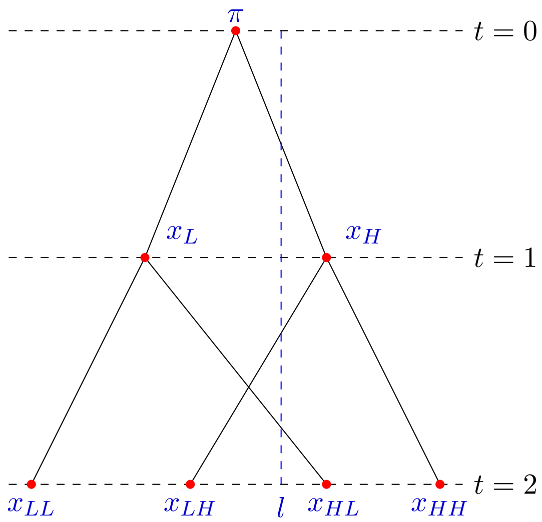

Figure 1: Two period model with binary signals.

Another interesting phenomenon in this simple example is the non-monotonic relationship between the optimality of the greedy policy and the precision of the second signal. If we set $q_2 = 1/2$ or $q_2 = 1$, the greedy policy is optimal for all w_2 in both cases. However, as shown above, this is not true for $q_2 = 4/5$.

Solving these optimization problems can be done by observing that the only free parameter of the optimization problem is the probability of adoption at time $t = 1$. Once this parameter is set, the receiver's belief conditional on adoption being l uniquely defines the probability of adoption at x_L and at x_H. At time $t = 2$, the policy is greedy with respect to the remaining mass. $\square$

[7] The numerical cutoff is approximate.

2.3 Measure Theoretic Formulation

The perspective of viewing persuasion as an optimal stopping problem is very useful conceptually, but it is not quite enough to help us solve the problem. Next we will provide a second perspective based on measures (and sub-measures) to reformulate our problem as a dynamic programming problem. This equivalent formulation relies on the kernel representation of Markovian martingales, with positive measures on $[0,1]$ as its state variables and sub-measures as its control.

A probability kernel is a measurable function $\sigma : [0,1] \to \Delta([0,1])$ such that $\mathbb{E}_{X\sim\sigma(x)}[X] = x$ for every $x \in [0,1]$. That is, the expectation under the probability measure $\sigma(x)$ is x for any $x \in [0,1]$. A probability kernel defines an operator from $\Delta([0,1])$ to itself where for every probability measure μ, $\sigma \circ \mu \in \Delta([0,1])$ is the *pushforward* probability measure of μ by σ such that for every Borel measurable set $B \subseteq [0,1]$ it holds that

$$\sigma \circ \mu\,(B) = \int_{[0,1]} \sigma(x)(B)\mathrm{d}\mu(x).$$

We note that a probability measure $\mu_1 \in \Delta([0,1])$ together with T probability kernels $\sigma_1, \ldots, \sigma_{T-1}$ determines a martingale $\mathbf{X} = (X_t)_{t=0,1,\ldots,T}$ where $X_1 \sim \mu_1$ and for every $1 \le t < T$

$$\mathbb{P}_{X_{t+1}}(\,\cdot\,|X_t = x_t, \ldots, X_1 = x_1) = \mathbb{P}_{X_{t+1}}(\,\cdot\,|X_t = x_t) = \sigma_t(x_t).$$

That is, the conditional distribution of X_{t+1} given $X_t = x_t, \ldots, X_1 = x_1$ is determined by $\sigma_t(x_t)$. We note that for every $t = 2, \ldots, T-1$ the distribution μ_t of X_t is given by $\sigma_{t-1} \circ \ldots \circ \sigma_1 \circ \mu_1$.

Conversely, for a given Markovian martingale $\mathbf{X} = (X_t)_{t=0,1,\ldots,T}$ supported on $[0,1]$, there exist probability kernels $\sigma_1, \ldots, \sigma_{T-1}$ such that if $X_1 \sim \mu_1$ then the martingale generated by the above procedure is $\mathbf{X} = (X_t)_{t=0,1,\ldots,T}$. This can be easily shown by taking $\sigma_t(x_t) = \mathbb{P}_{X_{t+1}}(\,\cdot\,|X_t = x_t)$. Note that for every $1 \le t \le T-1$ the kernel σ_t is a.s. uniquely defined on the support of X_t and can be arbitrarily defined outside of the support.

For any positive and finite measure ν over $[0,1]$ we let $|\nu| = \nu([0,1])$ and $\overline{\nu}$ be the expectation of its normalization $\frac{1}{|\nu|}\nu$. Given two positive measures ν, μ on $[0,1]$, we denote $\nu \le \mu$ if $\nu(B) \le \mu(B)$ for all Borel measurable sets $B \subseteq [0,1]$. Now we are ready to present an alternative formulation to the optimization problem.

Lemma 1. *The sender's optimization problem given in equation* (3) *has the fol-*

lowing equivalent reformulation:

$$\max_{\{\nu_t\}_{t=1}^T} \sum_{t=1}^{T} |\nu_t| w_t \tag{4}$$

subject to the following recursively defined constraints: $X_1 \sim \mu_1$, $\nu_t \leq \mu_t$ *for every* $t = 1, \dots, T$, $\overline{\nu_t} \geq l$ *for every* $t = 1, \dots, T$, *and* $\mu_t := \sigma_{t-1} \circ (\mu_{t-1} - \nu_{t-1})$ *for every* $t = 2, \dots, T$.

In this formulation, we have a mass μ_t that evolves according to the probability kernels σ_t, and in each step, we decide which parts of this mass, as described by ν_t, to eliminate from the Markov process to induce an immediate adoption. The proof of this lemma, as well as omitted proofs of other lemmas, are relegated to Appendix A.

3 Blackwell Order Preserving Kernels

In this section, we will introduce the concept of "Blackwell order preserving kernels," which proves to be fundamental to our analysis. Simply put, a probability kernel is Blackwell order preserving if a decision maker's preference for one probability measure of posteriors over another, in the Blackwell order, remains unchanged after receiving additional information represented by the kernel.

Recall that a probability measure of posteriors $\nu \in \Delta([0,1])$ dominates another probability measure $\mu \in \Delta([0,1])$ with respect to the Blackwell order [Blackwell, 1953], denoted as $\mu \preceq_B \nu$, if ν is a *mean preserving spread* of μ, or equivalently, if there exists a probability kernel σ such that $\sigma \circ \mu = \nu$.[8] We extend the Blackwell order to (finite) positive measures on $[0,1]$. For two positive measures μ, ν we write $\mu \preceq_B \nu$ if $|\mu| = |\nu|$ and $\frac{\mu}{|\mu|} \preceq_B \frac{\nu}{|\nu|}$. Now we are ready to introduce the central notion of "Blackwell order preserving kernels."

Definition 1. A probability kernel $\sigma : [0,1] \to \Delta([0,1])$ is called a *Blackwell order preserving* kernel if, for any two probability measures $\mu, \nu \in \Delta([0,1])$, $\mu \preceq_B \nu$ implies $\sigma \circ \mu \preceq_B \sigma \circ \nu$. A Markovian martingale is called *Blackwell order-preserving* if $\sigma_t : [0,1] \to \Delta([0,1])$ is Blackwell order-preserving for every time

[8] In terms of random variables one has $\mu \preceq_B \nu$ if and only if there exist random variables X and Y, where $X \sim \mu$, $Y \sim \nu$ and $\mathbb{E}[Y \mid X] = X$. In functional form, the order $\mu \preceq_B \nu$ is equivalent to requiring that $\nu(f) \leq \mu(f)$ for every non-decreasing concave $f : \mathbb{R} \to \mathbb{R}$, where we denote $\rho(f) = \int_{\mathbb{R}} f(x) \mathrm{d}\rho(x)$ for a positive measure ρ defined on $\mathbb{R}$.

$t = 1, \dots, T-1$.[9]

There is a simple interpretation of Blackwell order-preserving kernels. Consider a decision maker with a utility function $u : \Omega \times A \to \mathbb{R}$ who observes a signal about $\omega \in \Omega$ and then takes an action $a \in A$ to maximize their expected utility. By the definition of the Blackwell order [Blackwell, 1953], $\mu \preceq_B \nu$ if and only if the signal that induces the measure of posteriors ν generates a higher expected utility than the signal inducing the measure μ for *all* decision problems (i.e., all utility functions).

Now suppose that the decision-maker receives information in two periods where the second period's information is governed by a kernel σ based on the information received in the first period. Let ν and μ denote two possible distributions of the posterior in the first period, and $\sigma \circ \nu$ and $\sigma \circ \mu$ denote the two corresponding distributions of the posterior in the second period. The Blackwell order-preserving condition over σ simply requires that if all decision-makers prefer ν to μ in the first period then they will also prefer $\sigma \circ \nu$ and $\sigma \circ \mu$ in the second period. We note that a probability kernel preserves the Blackwell ordering between probability measures if and only if it preserves it between positive measures with equal mass.

The following lemma shows that, in order to verify that a kernel is Blackwell order preserving, it is sufficient (and necessary) to verify that the push-forward of every binary-supported distribution dominates the push-forward of its expectation in the Blackwell order.

Lemma 2. *A probability kernel $\sigma : [0,1] \to \Delta([0,1])$ is Blackwell order preserving if and only if, for every binary supported $\mu \in \Delta([0,1])$,*

$$\sigma \circ \delta_{\overline{\mu}} \preceq_B \sigma \circ \mu.$$

Similarly, a martingale $(X_t)_{t=1,\dots,T}$ is Blackwell preserving if and only if, for every $t \leq T-1$, the kernel σ_t satisfies the above condition for every binary supported measure μ such that the support of both μ and $\overline{\mu}$ is contained in the support of X_t.

The Blackwell order preserving property is a rather mild restriction on the belief martingales. We provide two classes of Blackwell order preserving martingales. The first class of martingales are belief martingales generated by **conditionally independent signals**. Formally, let S be some measurable signal space and let $G : \{0,1\} \to \Delta(S)$ be a probability kernel. We note that a prior $y \in [0,1]$ over the

[9]For martingales where σ_t is not uniquely defined, we only require that σ_t satisfies that there exists a version of σ_t that is Blackwell order preserving.

set $\Omega = \{0,1\}$ together with G generate a probability distribution $\mathbb{P}_y \in \Delta(\Omega \times S)$. Let $p_y(s) = \mathbb{P}_y(\omega = 1|s)$ be the conditional probability of $\{\omega = 1\}$ given s. Denote σ_y the posterior distribution of $p_y(s)$. That is, for every Borel subset $A \subseteq [0,1]$, $\sigma_y(A) = \mathbb{P}_y(p_y(s) \in A)$. Consider the probability kernel $\sigma : [0,1] \to \Delta(\{0,1\})$ defined by $\sigma : y \mapsto \sigma_y$. By the law of iterated expectation, this is indeed a probability kernel.[10]

Lemma 3. *Consider a martingale $\mathbf{X} = (X_t)_{t=1,\dots,T}$ that is generated by the kernels $(\sigma_t)_{t=1,\dots,T-1}$ such that for every $t = 1,\dots,T-1$ the kernel σ_t represent conditionally independent signal, then the martingale is a Blackwell order preserving martingale.*

The second class of martingales are belief martingales representing a **random walk**. Consider a discrete set $\Gamma = \{z_i\}_{i \in Z} \subseteq [0,1]$ such that $Z = \mathbb{Z} \cap [a,b]$ (where a and b might be finite or equal $-\infty$ and $+\infty$ respectively) $z_i > z_j$ for every $i > j \in Z$. Let σ be a kernel that represents a random walk on Γ. That is, $\sigma(z_i) = \delta_{z_i}$ if $i \in \{a,b\}$ and $\sigma(z_i) = \frac{z_{i+1}-z_i}{z_{i+1}-z_{i-1}}\delta_{z_{i-1}} + \frac{z_i - z_{i-1}}{z_{i+1}-z_{i-1}}\delta_{z_{i+1}}$, otherwise. A martingale $\mathbf{X} = (X_t)_{t=0,1,\dots,T}$ is a *random walk* if $X_0 = z_i$ with probability 1 for some $z_i \in \Gamma$ and $\sigma_t = \sigma$ for every $t = 1,\dots,T-1$. Random walk on the grid corresponds to the discrete version of the Brownian motion kernels considered in [Henry and Ottaviani, 2019, Orlov et al., 2020, Bizzotto et al., 2021, McClellan, 2022]. In Section 5, we will characterize the sender's optimal information policy when the sender's belief martingale is induced by a random walk.

Lemma 4. *A martingale that is induced by a random walk on a grid is Blackwell order-preserving.*

Remark 2. The Blackwell-order preserving property is an intuitive and rather mild restriction on belief martingales. Nevertheless, one can easily find a probability kernel that fails this property, if signals are allowed to be correlated across periods. Consider, for example, measures $\mu = \frac{1}{2}\delta_{\frac{1}{4}} + \frac{1}{2}\delta_{\frac{3}{4}}$ and $\nu = \delta_{\frac{1}{2}}$, and probability kernel σ with $\sigma\left(\frac{1}{2}\right) = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$, $\sigma\left(\frac{1}{4}\right) = \delta_{\frac{1}{4}}$, and $\sigma\left(\frac{3}{4}\right) = \delta_{\frac{3}{4}}$. It is clear that $\nu \preceq_B \mu$ but $\sigma \circ \mu \preceq_B \sigma \circ \nu$. In this example, the signal associated with σ is no longer conditionally independent of the signal that induces μ or ν. □

[10] In a different context, Kosenko [2021] showed that the garbling operator preserves the Blackwell ordering.

4 Main Characterization

In this section we introduce two classes of information policies, the "greedy policy" and the "interval policy." The optimality of the former policy is often the focus of the existing literature, see for example [Renault et al., 2017] and [Ely, 2017], while the latter policy is a generalization of the former and is crucial for our analysis. Our main characterization is that, if the sender's belief martingale is Blackwell-order preserving, then the sender's optimal strategy must be an interval policy.

4.1 The greedy policy

Consider the sender's optimization problem (3) with weights $w_1 = 1$ and $w_t = 0$ for $t \geq 2$. Consider a randomized stopping rule τ defined by functions τ_t as in (1), with $\tau_1(x_1) = 1$ if $x_1 \in (y, 1]$ and $\tau_1(x_1) = 0$ if $x_1 \in [0, y)$, where y is chosen to satisfy $\mathbb{E}[X_1|\tau = 1] = l$.[11] This stopping rule solves the sender's optimization problem because it maximizes the probability $\mathbb{P}(\tau = 1)$ across all stopping rules τ for which $\mathbb{E}[X_1|\tau = 1] \geq l$.

With $w_1 = 1$ and $w_t = 0$ for $t \geq 2$, the sender is myopic and *greedy*: she cares only about the persuasion probability at $t = 1$. The above function τ_1 solves the greedy sender's problem, and is thus referred to as the *greedy policy* with respect to μ_1. The measure $\nu_1 = \mathbb{P}(\,\cdot|\tau = 1)\,\mathbb{P}(\tau = 1)$ in the equivalent reformulation (4), which is eliminated from μ_1 by the above stopping rule, is referred to as the *greedy measure.*

The above greedy policy τ_1 is identified with the point y and the probability $\tau_1(y)$, which can be computed as follows. Note that a greedy measure can be obtained by taking the top q-quantile of the probability measure μ_1 such that the points above the quantile have a conditional expectation of l. More precisely, let F be the CDF of μ_1 and let $F^{-1}(x) = \inf\{x_1 | F(x_1) \geq x\}$. Then, for a $[0, 1]$-uniform random variable U, it holds that the random variable $\tilde{X}_1 = F^{-1}(U) \sim \mu_1$. The corresponding q-quantile is the unique value p such that $\mathbb{E}[\tilde{X}_1 | U \geq p] = l$. If μ_1 has no atoms, we define $y := F^{-1}(p)$ and allow $\tau_1(y)$ to take any values in $[0, 1]$. If $F^{-1}(p)$ is an atom of μ_1, we define $y := F^{-1}(p)$, and one can verify that $\tau_1(y)$ is given by

$$\tau_1(y) = \frac{F(F^{-1}(p)) - p}{\mu_1(\{F^{-1}(p)\})}.$$

[11] If μ_1 is discrete, randomization may occur at the cutoff point y.

For a martingale $\mathbf{X} = (X_t)_{t=1,\dots,T}$ with general weights, the greedy policy τ is defined recursively, where τ_1 is the greedy policy with respect to $X_1 \sim \mu_1$, and for every $t \geq 2$, τ_t is taken to be the greedy policy with respect to the measure $\mu_t = \mathbb{P}_{X_t}(\cdot|\tau \geq t)\,\mathbb{P}(\tau \geq t)$.

Definition 2. Let $\mathbf{X} = (X_t)_{t=1,\dots,T}$ be a martingale. A stopping rule τ is called a *greedy policy* if there exists a sequence of numbers $\{y_t\}_{t=1,\dots,T}$ with $y_t \in (0,1)$ such that for every $t = 1,\dots,T$, $\mathbb{E}[X_t|\tau = t] = l$ and

$$\tau_t(x_1,\dots,x_t) = \begin{cases} 1 & \text{if} \quad x_t \in (y_t, 1] \\ 0 & \text{if} \quad x_t \in [0, y_t). \end{cases}$$

4.2 The interval policy

Next, we define a class of policies that includes the greedy policy as a special case.

Definition 3. Let $\mathbf{X} = (X_t)_{t=1,\dots,T}$ be a martingale. A stopping rule τ is called an *interval policy* if there exists a sequence of intervals $\{[\underline{y}_t, \overline{y}_t]\}_{t=1,\dots,T}$ such that for every $t = 1,\dots,T$, $\mathbb{E}[X_t|\tau = t] = l$ and

$$\tau_t(x_1,\dots,x_t) = \begin{cases} 1 & \text{if} \quad x_t \in (\underline{y}_t, \overline{y}_t) \\ 0 & \text{if} \quad x_t \notin (\underline{y}_t, \overline{y}_t). \end{cases}$$

In terms of the mass elimination reformulation, one can induce an almost sure unique interval policy τ by choosing the eliminated sequence of masses $\{\nu_t\}_{t=1,\dots,T}$ such that $\nu_t(\underline{y}_t, \overline{y}_t) = \mu_t(\underline{y}_t, \overline{y}_t)$, $\nu_t([0,1] \setminus [\underline{y}_t, \overline{y}_t]) = 0$, and $\overline{\nu}_t = l$ for some sequence of intervals $\{[\underline{y}_t, \overline{y}_t]\}_{t=1,\dots,T}$. Therefore, an interval policy τ can be characterized by the interval $[\underline{y}_t, \overline{y}_t]$ at which it stops at time t and the probabilities, $\tau_t(\underline{y}_t)$ and $\tau_t(\overline{y}_t)$, that play a role only if μ_t has atoms on either $\underline{y}_t$ or $\overline{y}_t$. Generalizing the top quantile approach discussed for the greedy measure, we note that at each time t an interval policy may be identified with the two top quantiles $\{\underline{q}_t, \overline{q}_t\}$ of μ_t it induces. Formally, if we let F_{μ_t} be the CDF of μ_t, then $\underline{y}_t = F^{-1}_{\mu_t}(\underline{p}_t)$ and $\overline{y}_t = F^{-1}_{\mu_t}(\overline{p}_t)$, where $\mathbb{E}_{\tilde{X}_t \sim \mu_t}[\tilde{X}_t | U \in (\underline{p}_t, \overline{p}_t)] = l$, $\underline{q}_t = \mathbb{E}_{\tilde{X}_t \sim \mu_t}[\tilde{X}_t | U \geq \underline{p}_t]$, $\overline{q}_t = \mathbb{E}_{\tilde{X}_t \sim \mu_t}[\tilde{X}_t | U \geq \overline{p}_t]$,

$$\tau_t(\underline{y}_t) = \frac{F_{\mu_t}(F^{-1}_{\mu_t}(\underline{p}_t)) - \underline{p}_t}{\mu_t(\{F^{-1}_{\mu_t}(\underline{p}_t)\})},$$

and

$$\tau_t(\overline{y}_t) = \frac{F_{\mu_t}(F^{-1}_{\mu_t}(\overline{p}_t)) - \overline{p}_t}{\mu_t(\{F^{-1}_{\mu_t}(\overline{p}_t)\})}.$$

For any given measure μ_t in period t, there is a unique greedy policy and hence the persuasion probability. In contrast, there are many interval policies and different interval policies may have different persuasion probabilities. For any given persuasion probability (smaller than the one of the greedy policy), however, there is a unique interval policy that attains this probability, which allows us to view an interval policy as one dimensional object.

Lemma 5. *Let α be the persuasion probability of the greedy policy with respect to μ_1. For any $0 < \beta \leq \alpha$ there exists an almost surely unique interval policy τ for which $\mathbb{P}(\tau = 1) = \beta$.*

The interval policy discloses "minimal information" necessary for adoption in the sense that it minimizes the amount of information being disclosed among all policies that attains a given persuasion probability in a given period.[12] To formally establish this, recall that for two positive measures μ, ν on $[0,1]$ with $|\mu| = |\nu|$, we can extend the Blackwell ordering and write $\mu \preceq_B \nu$ if $\frac{\mu}{|\mu|} \preceq_B \frac{\nu}{|\nu|}$. The Blackwell ordering relation forms a partial order in the class of all stopping rules $\{\nu : \nu \leq \mu_1\}$ that yields a given persuasion probability. The following lemma establishes that the interval policy is the minimal element with respect to this partial order.

Lemma 6. *Let $\nu' \leq \mu_1$ be the measure that corresponds to the interval policy that stops on $[\underline{y}, \overline{y}]$ with $\overline{\nu'} = l$. Let ν'' with $\nu'' \leq \mu_1$ and $\overline{\nu''} = l$ be another stopping rule that satisfies $|\nu''| = |\nu'|$. Then $\nu' \preceq_B \nu''$.*

4.3 Optimality of interval policies

Now we are ready to state our main characterization of optimal information policies. It demonstrates a strong connection between interval policies and Blackwell order-preserving martingales.

Theorem 1. *If $\mathbf{X} = (X_t)_{t=1,\ldots,T}$ is a Blackwell order preserving martingale, then there exists an interval policy that solves the sender's optimization problem (3).*

This characterization removes a major hurdle in solving the sender's optimization problem. In the equivalent formulation (4) of the sender's optimization problem (3), the choice variables are a sequence of measures $\{\nu_j : j \geq 1\}$ which

[12] This is connected to but different from the "minimal information" property of the greedy policy in [Renault et al., 2017]. Their greedy policy reveals nothing ("minimal information") if the receiver's current belief lies in the region that supports the sender's preferred action; otherwise discloses to maximize the persuasion probability in the current period.

are infinite dimensional. By Theorem 1, it is without loss to restrict attention to interval policies. By Lemma 5, each interval policy can be identified with a sequence of one-dimensional probability mass $\{|\nu_j| : j \geq 1\}$ it eliminates in each period. Therefore, the sender only needs to decide the size of probability mass in each period she aims to persuade the receiver at that period. Once this decision is made, the sequence of "interval" measures $\{\nu_j : j \geq 1\}$ is defined uniquely.

Theorem 1 does not characterize the optimal sequence of stopping masses $\{|\nu_j| : j \geq 1\}$. As shown in Example 1 (and Examples 2 below), these stopping masses will depend on the sender's payoff functions or the sequence of weights $\{w_j : j \geq 1\}$. Instead, it uniquely specifies the optimal policy as a function of these stopping masses. Furthermore, Theorem 1 implies that an optimal interval policy exists for the case $T = \infty$. Thus the "sup" in the optimization problem (3) can be replaced with "max" for Blackwell order-preserving martingales.

Remark 3. As noted in Lemma 6, the interval policy is minimal with respect to the Blackwell order among all stopping rules with the same size of the stopping probability mass. This means that the interval policy reveals the least amount of information necessary to achieve a given persuasion probability. Consequently, it retains the maximum amount of information at the end of the current period if the game continues. The Blackwell order preserving property ensures that this information advantage is maintained at the beginning of the next period. This is essentially the idea of the proof for Theorem 1. □

Before proving Theorem 1, we first uses an example to illustrate how our main characterization can be used to solve the optimal information policy. Then we present another example to show why the Blackwell order preserving property is necessary for our characterization.

Example 2 (Two Periods, Continuous Signals)**.** We demonstrate here how to utilize Theorem 1 in an example with $T = 2$. We set $w_1 = 1$. Suppose that the sender receives two conditionally independent signals, $S_1 \in [0, 1]$ in period 1 and $S_2 \in \{0, 1, \phi\}$ in period 2. The common prior over the binary state is $\pi = 1/2$. The distribution of the first period signal S_1, conditional on the state, is given by

$$\mathbb{P}(S_1 \leq s|\omega = 1) = s^2 \quad \text{and} \quad \mathbb{P}(S_1 \leq s|\omega = 0) = s(2 - s).$$

It is easy to verify that signal S_1 induces a uniform posterior over $[0, 1]$. The second period signal S_2 is either a perfect one or a white noise: with probability $q \in [0, 1]$, signal S_2 perfectly reveals the state $\omega \in \{0, 1\}$, and with probability $1 - q$, S_2 is a null signal (denoted by ϕ) which does not contain any information.

It is easy to verify that the martingale $\mathbf{X} = (X_1, X_2)$ must satisfy $X_1(S_1) = S_1$ and

$$X_2(S_1, S_2) = \begin{cases} 1 & \text{if} \quad S_2 = 1 \\ 0 & \text{if} \quad S_2 = 0 \\ S_1 & \text{if} \quad S_2 = \phi \end{cases}$$

Therefore, X_1 is uniformly distributed on $[0, 1]$, and X_2 is also uniformly distributed but with atoms at 0 and 1 splitting a total mass of q, as in Figure 2.

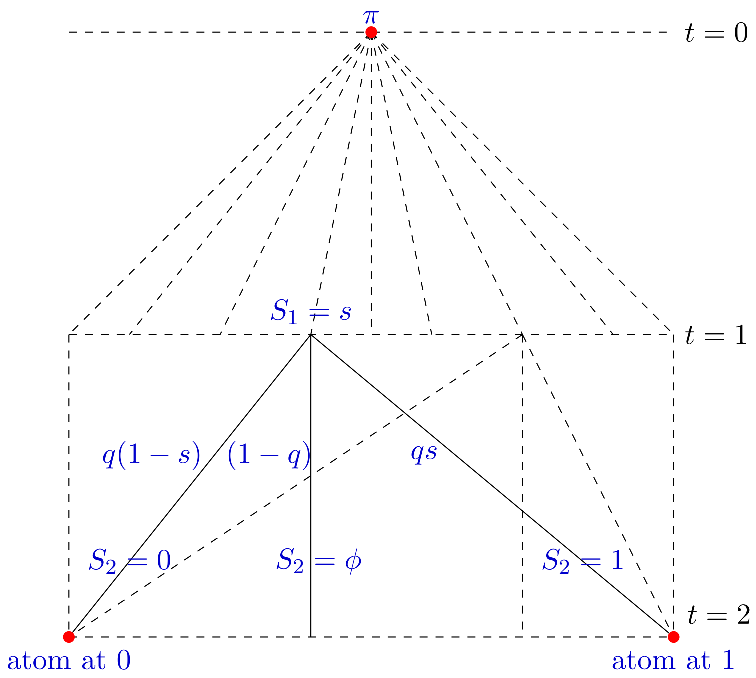

Figure 2: Two period model with continuous signals

What is the optimal persuasion policy in this example? Suppose that in the optimal policy a probability mass of size $\alpha \leq 2(1-l)$ is persuaded in period 1. That is, $|\nu_1| = \alpha$ and $\overline{\nu_1} = l$. By Theorem 1, this probability mass must be taken from the interval $[l - \alpha/2, l + \alpha/2]$. Let $\hat{x}(\alpha)$ denote the optimal cutoff such that all probability mass remaining in period 2 with $X_2 \geq \hat{x}(\alpha)$ will be persuaded in period 2. With some algebra, one can show that $\hat{x}(\alpha)$ is given by

$$\hat{x}(\alpha) = l - \sqrt{\frac{(1-l)^2 + (1-l)lq(1-2\alpha)}{1-q}}$$

Again we normalize the sender's utility by setting $w_1 = 1$. Then we can write the sender's utility as a function of α:

$$\Gamma(\alpha) = \alpha + w_2 \left[(1-q)(1 - \alpha - \hat{x}(\alpha)) + \frac{1}{2} q(1 - 2l\alpha) \right],$$

where the first term in the bracket captures the uniform probability mass that is truncated below by $\hat{x}(\alpha)$ and is persuaded when $S_2 = \phi$, and the second term represents the atom at $X_2 = 1$ formed when $S_2 = 1$. Both terms take into account the fact that the probability mass of α is persuaded in period 1. It is easy to verify that $\Gamma(\alpha)$ is concave in α for all feasible $\alpha \in [0, 2(1-l)]$.

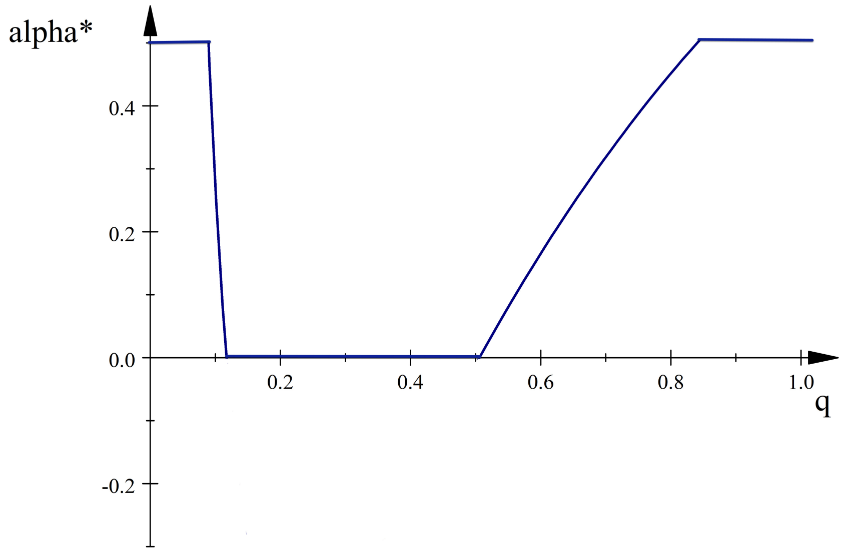

Figure 3: Optimal α^* as a function of the signal precision q in period two.

The optimal policy is characterized a pair of cutoff functions, $w_L(q,l)$ and $w_H(q,l)$ with $w_L(q,l) < w_H(q,l)$. If $w_2 \le w_L(q,l)$, $\Gamma(\alpha)$ is increasing for all $\alpha \in [0, 2(1-l)]$ and hence the optimal policy is greedy in both periods with $\alpha^* = 2(1-l)$. If $w_2 \ge w_H(q,l)$, $\Gamma(\alpha)$ is decreasing for all $\alpha \in [0, 2(1-l)]$ and hence the optimal policy persuades only in period 2 with $\alpha^* = 0$. Finally, if $w_2 \in (w_L(q,l), w_H(q,l))$, an interior $\alpha^* \in (0, 2(1-l))$ is optimal and varies with q. Interestingly, for a typical fixed l, as q increases from 0, optimal α^* is first $2(1-l)$, then interior, then 0, then interior again, and finally $2(1-l)$ again as q approaches 1, as shown in the Figure 3 (where we take $l = 3/4, w_2 = 24/25$). $\square$

In our next example, the probability kernel is not Blackwell-order preserving and the optimal policy is not an interval policy. Therefore, the Blackwell order preserving property is necessary for Theorem 1.

Example 3 (Interval Policy Is Sub-optimal)**.** Consider a two-period model with weights $w_1 = 1$ and $w_2 = \frac{3}{4}$. Let $\mu_1 = \frac{1}{7}\delta_{\frac{1}{3}} + \frac{2}{7}\delta_{\frac{1}{2}} + \frac{4}{7}\delta_{\frac{3}{4}}$ and $l = \frac{2}{3}$. The kernel σ is defined as follows:

$$\sigma\left(\frac{1}{2}\right) = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_0, \quad \sigma\left(\frac{1}{3}\right) = \delta_{\frac{1}{3}}, \quad \text{and} \quad \sigma\left(\frac{3}{4}\right) = \delta_{\frac{3}{4}}.$$

As we argued in Remark 2, this kernel is not Blackwell order preserving. The greedy measure in period one is $\nu_1 = \frac{2}{7}\delta_{\frac{1}{2}} + \frac{4}{7}\delta_{\frac{3}{4}}$ with $\overline{\nu}_1 = l$. Therefore, a policy ν_1' is an interval policy if and only if $\nu_1' := \nu_1^\beta = (1-\beta)\nu_1$ for some $\beta \in [0,1]$. The period-two measure after taking out probability measure ν_1' in period one is

$$\mu_2^\beta := \sigma \circ (\mu_1 - (1-\beta)\nu_1) = \beta\frac{1}{7}\delta_0 + \frac{1}{7}\delta_{\frac{1}{3}} + \beta\frac{4}{7}\delta_{\frac{3}{4}} + \beta\frac{1}{7}\delta_1.$$

Applying the greedy policy to μ_2^β yields the greedy measure in period two:

$$\nu_2^\beta = \begin{cases} \beta(\frac{2}{7}\delta_{\frac{1}{3}} + \frac{4}{7}\delta_{\frac{3}{4}} + \frac{1}{7}\delta_1) & \text{if} \quad \beta \le \frac{1}{2} \\ \left(\beta - \frac{1}{2}\right)\frac{1}{7}\delta_0 + \frac{1}{7}\delta_{\frac{1}{3}} + \beta\frac{4}{7}\delta_{\frac{3}{4}} + \beta\frac{1}{7}\delta_1) & \text{if} \quad \beta > \frac{1}{2} \end{cases}$$

Overall applying $(\nu_1^\beta, \nu_2^\beta)$ yields a utility:

$$w_1 \cdot |\nu_1^\beta| + w_2 \cdot |\nu_2^\beta| = \begin{cases} (1-\beta)\frac{6}{7} + \frac{3}{4}\beta & \text{if} \quad \beta \le \frac{1}{2} \\ (1-\beta)\frac{6}{7} + \frac{3}{4}\left(\frac{6}{7}\beta + \frac{1}{14}\right) & \text{if} \quad \beta > \frac{1}{2} \end{cases}$$

Therefore, $\beta = 0$ is optimal and the utility for the optimal interval policy is $\frac{6}{7}$.

Consider an alternative policy of $\tilde{\nu}_1 = \frac{1}{7}\delta_{\frac{1}{3}} + \frac{4}{7}\delta_{\frac{3}{4}}$. This leaves the measure in period two as $\tilde{\mu}_2 := \sigma \circ (\mu_1 - \tilde{\nu}_1) = \frac{1}{7}\delta_0 + \frac{1}{7}\delta_1$ with the greedy measure $\tilde{\nu}_2 = \frac{1}{14}\delta_0 + \frac{1}{7}\delta_1$. The utility of this policy is

$$\frac{5}{7} + \frac{3}{4} \cdot \frac{3}{14} = \frac{7}{8} > \frac{6}{7}.$$

This implies that any interval policy is sub-optimal for this problem. $\square$

4.4 Proof of Theorem 1

We start the proof by establishing the following simple claim. For any two measures with $\nu \le \mu_1, \nu' \le \mu_1$, $\overline{\nu} = \overline{\nu'}$ and $|\nu| = |\nu'|$, it holds that

$$\nu \preceq_B \nu' \implies \mu_1 - \nu' \preceq_B \mu_1 - \nu. \tag{5}$$

This is immediate because, for every $x \in [0,1]$,

$$\int_0^x \nu([0,y])dy \ge \int_0^x \nu'([0,y])dy \implies \int_0^x (\mu_1 - \nu)([0,y])dy \le \int_0^x (\mu_1 - \nu')([0,y])dy.$$

Now consider any sender strategy $\{\nu_t', \mu_t'\}_{t=1,\dots,T}$ that satisfies the recursive formulation of Lemma 1. We will show by an induction argument that there exists an interval policy $\{\nu_t, \mu_t\}_{t=1,\dots,T}$ that has $|\nu_t| = |\nu_t'|$ for $t = 1, \dots, T$ and also satisfies the conditions of Lemma 1.

Specifically, we will prove by induction the hypothesis that for every $t \leq T$, if $\{\nu_j, \mu_j\}_{j=1,\ldots,t-1}$ satisfy $|\nu_j| = |\nu'_j|$ for all $j \leq t-1$ and $\mu'_t \preceq_B \mu_t$, then there exists an interval policy ν_t with $|\nu_t| = |\nu'_t|$ such that $\nu_t \preceq_B \mu_t$ and $\mu'_{t+1} \preceq_B \mu_{t+1}$. The proof will repeatedly apply Lemmas 5 and 6, and claim (5).

For $t = 1$, the hypothesis follows from Lemma 5 that there exists an interval measure ν_1 with $|\nu_1| = |\nu'_1|$ and $\nu_1 \leq \mu_1$. Lemma 6 implies that $\nu_1 \preceq_B \nu'_1$, and hence $\mu'_1 - \nu'_1 \preceq_B \mu_1 - \nu_1$ by claim (5) (note that $\mu_1 = \mu'_1$). Since σ_1 is Blackwell order preserving it follows that $\mu'_2 = \sigma_1 \circ (\mu'_1 - \nu'_1) \preceq_B \sigma_1 \circ (\mu_1 - \nu_1) = \mu_2$ as desired.

Now suppose the hypothesis holds for $t - 1 \leq T - 1$, and we want to prove it for t. Since $\mu'_t \preceq_B \mu_t$ by the induction hypothesis, there exists a probability kernel $\kappa : [0,1] \to \Delta([0,1])$ such that $\kappa \circ \mu'_t = \mu_t$. Let $\psi_t = \kappa \circ \nu'_t$. It holds that $|\psi_t| = |\nu'_t|$ and $\overline{\psi}_t = \overline{\nu'}_t = l$. Since $\mu_t - \psi_t = \kappa \circ (\mu'_t - \nu'_t)$ it holds that

$$\mu'_t - \nu'_t \preceq_B \mu_t - \psi_t. \tag{6}$$

By Lemmas 5 and 6, there exists an interval policy $\nu_t \leq \mu_t$ such that $|\nu_t| = |\nu'_t| = |\psi_t|$, $\nu_t \preceq_B \psi_t$, and $\overline{\nu}_t = l$. It follows from claim (5) that

$$\mu_t - \psi_t \preceq_B \mu_t - \nu_t. \tag{7}$$

Conditions (6) and (7), together with the fact that σ_t is Blackwell order preserving, yield that

$$\mu'_{t+1} = \sigma_t \circ (\mu'_t - \nu'_t) \preceq_B \sigma_t \circ (\mu_t - \nu_t) = \mu_{t+1},$$

which concludes the induction step.

Therefore, we have shown that for every feasible policy there exists an internal policy that achieves at least the same utility for the sender. This implies that the "sup" in the sender's problem (3) can be taken over interval policies to attain the optimal utility for the sender. Proposition 2 in Appendix C shows that an optimal interval policy always exists (even for $T = \infty$), and therefore the "sup" can be further replaced by "max".

5 Application to Random Walks

In this section, we assume that the sender's belief martingale follows a random walk on a discrete grid, introduced in Section 3. Recall that the set of grid points is $\Gamma = \{z_j\}_{j \in Z}$ with $Z = \mathbb{Z} \cap [a, b]$ and $z_j < z_{j+1}$ for any $j \in Z$. It is without loss (by relabeling the indexes) to assume that $l \in (z_0, z_1]$. We define

$$\overline{\delta} = \overline{\delta}(\Gamma) := \inf_{j \leq 0} \frac{l - z_j}{l - z_{j-1}}. \tag{8}$$

Namely, $\overline{\delta} \leq 1$ measures the "maximal scale" by which the random walk may *shrink* the distance from the threshold l in a single upward jump from z_{j-1} to z_j, when the ending belief z_j remains strictly below l.

We further assume that $w_t = \delta^{t-1}$ for some discount factor $\delta > 0$ and the common prior is distributed as the Dirac measure on z_j for some grid point $z_j \in \Gamma$ (i.e., $X_0 \sim \delta_{z_j}$). The horizon T can be either finite or infinite.

Theorem 2. *If the sender's belief martingale is a random walk on the grid Γ and she discounts future with $\delta \leq \overline{\delta}(\Gamma)$, then the greedy policy is optimal.*

The theorem states that for a sufficiently impatient sender (i.e., with a discount factor below $\overline{\delta}$), the greedy policy is optimal. We first discuss below the tight connection between the greedy policy and transparency. We then calculate $\overline{\delta}$ for several examples in Section 5.1. In Section 5.3 we use an example to demonstrate that the greedy policy can be sub-optimal for a sufficiently large discount factor with $T = \infty$. Section 5.2 discusses the ideas of the proof for Theorem 2, while the formal proof appears in Section 5.4.

As the greedy policy is defined (see Section 4.1), the sender stays mute until she makes an "adopt" recommendation, while by transparency we mean the opposite extreme in which the sender reveals her private information in each period. Despite these seemingly two opposite extremes, the arguments below show that, when the martingale is a random walk, there exists a policy that is equivalent (in terms of adoption) to the greedy policy and is almost fully transparent.

Consider the policy that fully reveals the posterior x_t (i.e., reveals the entire private information) as long as $x_t = z_j$ for $j \leq 0$. Once her posterior reaches $x_t = z_0$, at time $t+1$ she either reveals that $x_{t+1} = z_{-1}$ (with a weakly lower probability than it actually happens) or, alternatively, sends an "adopt" recommendation to induce the posterior l. It is easy to see that the event of adoption under this policy is identical to the event of adoption under the greedy policy. But unlike the greedy policy, here the sender stays fully transparent about her private information in all periods until it becomes very close to the threshold (i.e., at the moment her posterior belief reaches $x_t = z_0$).

5.1 Examples: Compute $\overline{\delta}(\Gamma)$ for different grid Γ

We demonstrate how the threshold $\overline{\delta}(\Gamma)$ in Theorem 2 can be computed in several examples. In all these examples, the infimum in the definition (8) is often attained at $j = 0$. That is,

$$\overline{\delta} = \frac{l - z_0}{l - z_{-1}}.$$

Standard grid. For every $\epsilon > 0$ let $\Gamma = \{n\epsilon : n \in \mathbb{N}_0, n\epsilon \leq 1\}$ be the ϵ-grid, where $\mathbb{N}_0$ denotes the set of natural numbers including 0. If the threshold is located on the grid ($l \in \Gamma$), then $\overline{\delta} = 1/2$ and the minimum is obtained at $j = 0$. However, if $l \notin \Gamma$, say $l = n\epsilon + \epsilon'$ for $\epsilon' < \epsilon$, the bound is given by $\overline{\delta} = \frac{\epsilon'}{\epsilon'+\epsilon}$, which is smaller than 1/2 and again obtained at $j = 0$.

Conditionally i.i.d. binary symmetric signals. Consider a scenario in which the sender observes in each period a signal that matches the state ω with probability $p \in (\frac{1}{2}, 1)$, independently across periods (conditional on the state). For simplicity, assume that $\frac{1}{2} \in \Gamma$. In this case, Γ takes the following simple form:

$$\Gamma = \left\{ \frac{p^z}{p^z + (1-p)^z} : z \in \mathbb{Z} \right\}.$$

Assume the prior $\pi < \frac{1}{2}$ and $l = z_1 = \frac{1}{2}$. Again, the infimum in (8) is obtained at $j = 0$ and

$$\overline{\delta} = \frac{\frac{1}{2} - \frac{p^{-1}}{p^{-1}+(1-p)^{-1}}}{\frac{1}{2} - \frac{p^{-2}}{p^{-2}+(1-p)^{-2}}} = 2p^2 - 2p + 1.$$

As the signals of the sender become more accurate ($p \to 1$), the bound $\overline{\delta} \to 1$, and the greedy policy becomes optimal for an arbitrary patient sender.

A standard grid with a hole. The greedy policy is optimal for an arbitrary patient sender under another class of grids. Start with the standard grid but eliminate grid points that are close to and below l. Formally, let $\epsilon' \gg \epsilon$ and $\Gamma = \{n\epsilon : n \in \mathbb{N}_0, n\epsilon \leq 1, n\epsilon \notin (l - \epsilon', l)\}$. Namely, from the point $x_t = l - \epsilon'$ the belief of the sender either jumps down to the nearby point $x_{t+1} = l - \epsilon' - \epsilon$ or jumps up to the far point (weakly) above l. In this case, $\overline{\delta} = \frac{\epsilon'}{\epsilon'+\epsilon}$, which approaches 1 as $\epsilon/\epsilon' \to 0$.

5.2 Idea of the proof of Theorem 2

We start with the case $T < \infty$ and prove the theorem using backward induction. By the single deviation principle (or equivalently the backward induction hypothesis), it is sufficient to prove that, for every state $X_0 \in \Delta(\Gamma)$, if the sender acts greedily starting from tomorrow, she will be better off acting greedily today as well.

Theorem 1 allows us to restrict the set of possible policies significantly: the decision in each state X_1 is characterized by an interval $[\underline{y}, \overline{y}]$ such that the conditional mean of X_1 over $[\underline{y}, \overline{y}]$ is l. Namely, there is no need to consider the set of all sub-measures whose mean is l but only those that are supported on an interval.

The proof that acting greedily today is indeed superior is conducted in two steps. First, we demonstrate that the utility of acting greedily in the first two days provides the best possible utility for the first two periods among all possible interval policies (see Lemma 11 and Corollary 2). Here we rely on the assumption that $\delta \leq \overline{\delta}$.[13]

In the second step, we prove that the remaining mass after two consecutive applications of the greedy action is superior to the remaining mass after any interval action followed by a single greedy action (see Lemma 7). Here, we shall clarify the sense in which it is superior. The second-order stochastic dominance partial order is irrelevant because we compare two measures with different masses. Although the first-order stochastic dominance (FOSD) is satisfied, there is a complication: the FOSD order is not necessarily preserved under a random walk over a grid. Therefore, a weaker (than FOSD) version of dominance is required for our arguments (see Definition 4), which relates to the notion known in the literature as the *increasing convex order* (see, e.g., [Shaked and Shanthikumar, 2007]). Our notion of domination seems to extend this concept to general positive finite measures.

This order exactly serves our purpose. On one hand, it is preserved by a random walk on a grid (see Lemma 9). On the other hand, it is sufficiently powerful to deduce that the sender will be better off by remaining with the dominant measure (see Lemma 10).

To summarize, we show that the greedy policy is superior to any other interval policy in both aspects: it provides a better utility in the first two periods (the first step above) and it leaves the sender with a measure that she can better utilize in the future periods (the second step above).

A closer examination of our proof reveals that it does not rely entirely on the

[13]Without this assumption, this claim is false as shown in Proposition 1 below.

assumption that the sender is initially uninformed (i.e., $X_0 = \delta_\pi$). In fact, the argument based on backward induction holds more generally. In Section 5.4, we define the set of measures $\Delta^*(\Gamma)$, which consists of all positive measures supported on either the even or odd points of the grid.

Our proof that the greedy policy is optimal remains valid under the weaker assumption that $X_0 \sim \mu_0 \in \Delta^*(\Gamma)$ with $\mathbb{E}[X_0] < l$. Thus the proof of Theorem 2 implies the following corollary.

Corollary 1. *Suppose that the sender's belief martingale follows a random walk on the grid Γ and she is initially partially informed with $X_0 \sim \Delta^*(\Gamma)$. If her discount factor satisfies $\delta \leq \overline{\delta}(\Gamma)$, then the greedy policy is optimal.*

5.3 Sub-optimality of the greedy policy for a patient sender

It remains an open question whether the greedy policy is optimal for an initially uninformed sender, whose belief martingale follows a random walk on a grid (either the standard grid or one induced by conditionally i.i.d. binary signals). In this subsection, however, we demonstrate that Corollary 1 fails if the sender is sufficiently patient. This implies that, for the greedy policy to be optimal for any initially partially informed sender, some constraints on the sender's patience are necessary. This observation highlights the challenges in addressing the aforementioned open problem. The single-deviation principle (or the Bellman equation) is a central tool for proving results in such dynamic settings. Our negative finding suggests that alternative methods will be required to prove the optimality of the greedy policy, if it is indeed optimal.

Proposition 1. *Suppose $T = \infty$ and consider the standard ϵ-grid with $\Gamma = \{n\epsilon : n \in \mathbb{N}_0, n\epsilon < 1\}$. For every $\delta > \sqrt{2}/2$, there exists ϵ' such that, for all ϵ-grid Γ with $l \in \Gamma$ and $\epsilon < \epsilon'$, there is an initial prior $X_0 \in \Delta(\Gamma)$ for which the greedy policy is sub-optimal.*

The formal proof is relegated to Appendix B. It consists of several steps. We first assume by contradiction that the greedy policy is always optimal and denote by $v(j\epsilon)$ the sender's value from the prior $\delta_{j\epsilon}$ for any grid point $j\epsilon$. Using coupling considerations, we show that the optimality of the greedy policy implies that $v(l-2\epsilon) \approx (v(l-\epsilon))^2$, which is then used to approximate $v(l-\epsilon)$ as a function of δ and ϵ.

We then consider a prior that is supported on the points $l-2\epsilon$ and 1, and has an expectation below l. Since using the greedy policy leaves, at the next stage, only mass on the point $l-2\epsilon$, we use the above approximation to estimate the value of the greedy policy. We complete the proof by showing that the sender can achieve a higher utility than the greedy policy by waiting for a single period without using any mass and then pooling the mass obtained at $l-\epsilon$ together with the mass at 1.

In light of Proposition 1 an interesting question that we leave open is whether the bound in Theorem 2 is tight.

5.4 Proof of Theorem 2

Recall that given two positive measures ϕ, μ on $[0,1]$ we denote $\phi \leq \mu$ if $\phi(B) \leq \mu(B)$ for any Borel measurable set $B \subseteq [0,1]$. We now introduce relations between positive measures on $[0,1]$ that play a fundamental role in our analysis. The first is the well-known first order stochastic domination. A measure μ first order stochastic dominates (FOSD) a measure λ ($\lambda \preceq_F \mu$) if $\mu([x,1]) \geq \lambda([x,1])$ for any $x \in [0,1]$. A function $\rho : [0,1] \to \Delta([0,1])$ is called a FOSD kernel if $\varphi(x)([x,1]) = 1$. It is easy to see that $\lambda \preceq_F \mu$ iff there exists a FOSD kernel φ and a measure $\phi \leq \mu$ such that $\varphi \circ \lambda = \phi$.

We first introduce a central notion for our analysis which we call *domination*.

Definition 4. Say that a measure μ *dominates* a measure λ ($\lambda \preceq_D \mu$) if there exists a FOSD kernel $\varphi : [0,1] \to \Delta([0,1])$ and a probability kernel $\rho : [0,1] \to \Delta([0,1])$ such that there exists a measure $\phi \leq \mu$ such that $\rho \circ \varphi \circ \lambda = \phi$.

It follows directly from the definition that the notion of domination extends first and second-order stochastic domination. Namely, $\lambda \preceq_F \mu \Rightarrow \lambda \preceq_D \mu$ and $\lambda \preceq_B \mu \Rightarrow \lambda \preceq_D \mu$. In the special case where $|\lambda| = |\mu|$, the domination order is called in the *increasing convex order* in the literature; see [Shaked and Shanthikumar, 2007].[14]

We start by showing that two consecutive applications of the greedy action leaves a better probability mass than any interval action followed by a single greedy action.

Lemma 7. *Consider an infinite Blackwell-order preserving martingale $\mathbf{X} = (X_t)_{t\geq 1}$. Consider two interval policies (ν_1, ν_2) and (ν_1', ν_2') for the two first periods such that (ν_1, ν_2) is the greedy policy and (ν_1', ν_2') is any other policy where ν_2' is greedy. Then $\mu_2' - \nu_2' \preceq_F \mu_2 - \nu_2$.*

[14] We omit the proof of the observation that our notion of domination is equivalent to the increasing convex order because our proof does not rely on this observation.

Proof. By definition, $\mu_2 = \sigma_1 \circ (\mu_1 - \nu_1)$ and $\mu_2' = \sigma_1 \circ (\mu_1 - \nu_1')$. Since ν_1 is greedy and ν_1' is an interval policy, $\nu_1' \leq \nu_1$ and thus we can write $\nu_1 = \nu_1' + \psi$ for some positive measure ψ. Furthermore, since $\overline{\nu}_1 = \overline{\nu_1'} = l$, we have $\overline{\psi} = l$. It follows that

$$\mu_2 = \sigma_1 \circ (\mu_1 - \nu_1) = \sigma_1 \circ (\mu_1 - \nu_1' - \psi) = \mu_2' - \sigma_1 \circ \psi,$$

Therefore, there is a positive measure $\phi := \sigma_1 \circ \psi$ with $\overline{\phi} = l$ such that

$$\mu_2' = \mu_2 + \phi. \tag{9}$$

Next, by introducing $\lambda = \mu_2 - \nu_2$ we may further decompose μ_2' to

$$\mu_2' = \lambda + \nu_2 + \phi. \tag{10}$$

As $\overline{\nu_2} = l$ and $\overline{\phi} = l$, by the linearity of the integral operator, we have $\overline{\nu_2 + \phi} = l$ as well. Therefore, using decomposition (10), we infer that $\nu_2 + \phi$ defines a stopping rule with respect to μ_2'. As ν_2' is greedy with respect to μ_2', we have $\nu_2'([x, 1]) \geq (\nu_2 + \phi)([x, 1])$ for every $x \in [0, 1]$, because a greedy stopping rule first-order stochastic dominates all other obedient stopping rules. Thus, for every $x \in [0, 1]$,

$$\begin{aligned}
(\mu_2' - \nu_2')([x, 1]) &= \mu_2'([x, 1]) - \nu_2'([x, 1]) \\
&\leq \mu_2'([x, 1]) - (\nu_2 + \phi)([x, 1]) \\
&= (\mu_2 - \nu_2)([x, 1]),
\end{aligned}$$

where the last equality follows from (9). This indeed shows that $\mu_2 - \nu_2$ first-order stochastically dominates $\mu_2' - \nu_2'$, as desired. □

The domination relation defined above is closed under addition, as is established in the following lemma.

Lemma 8. *Assume that $\psi \preceq_D \mu$ and that $\psi' \preceq_D \mu'$ for some positive measure μ, μ', ψ, ψ' on a discrete grid Γ. Then it holds that $\psi + \psi' \preceq_D \mu + \mu'$.*

Proof. By definition, there exists φ, ρ and φ', ρ' such that $\rho \circ \varphi \circ \psi := \phi \leq \mu$ and $\rho' \circ \varphi' \circ \psi' := \phi' \leq \mu'$. Define a kernel $\tilde{\varphi}$ as follows:

$$\tilde{\varphi}(x) = \begin{cases} \frac{\psi(x)}{(\psi+\psi')(x)}\varphi(x) + \frac{\psi'(x)}{(\psi+\psi')(x)}\varphi'(x) & \text{if} \quad (\psi + \psi')x) > 0 \\ \delta_x & \text{if} \quad (\psi + \psi')(x) = 0 \end{cases}$$

Both ψ and ψ' are FOSD kernels, so $\tilde{\varphi}$ is also a FOSD kernel. Moreover,

$$\begin{aligned}\tilde{\varphi} \circ (\psi + \psi')(y) &= \sum_{x \in \Gamma,\ (\psi+\psi')(x) > 0} (\psi + \psi')(x) \tilde{\varphi}(x)(y) \\ &= \sum_{x \in \Gamma,\ (\psi+\psi')(x) > 0} (\psi + \psi')(x) \Big(\frac{\psi(x)}{(\psi + \psi')(x)} \varphi(x)(y) + \frac{\psi'(x)}{(\psi + \psi')(x)} \varphi'(x)(y) \Big) \\ &= \sum_{x \in \Gamma,\ (\psi+\psi')(x) > 0} \psi(x)\varphi(x)(y) + \psi'(x)\varphi'(x)(y) \\ &= \varphi \circ \psi(y) + \varphi' \circ \psi'(y).\end{aligned}$$

Therefore, $\tilde{\varphi}$ is a FOSD kernel and satisfies $\tilde{\varphi} \circ (\psi + \psi') = \varphi \circ \psi + \varphi' \circ \psi'$.

Next, similarly define a probability kernel $\tilde{\rho}$ as

$$\tilde{\rho}(x) = \begin{cases} \frac{\varphi\circ\psi(x)}{(\varphi\circ\psi + \varphi'\circ\psi')(x)} \rho(x) + \frac{\varphi'\circ\psi'(x)}{(\varphi\circ\psi + \varphi'\circ\psi')(x)} \rho'(x) & \text{if } \ (\varphi \circ \psi + \varphi' \circ \psi')(x) > 0 \\ \delta_x & \text{if } \ (\varphi \circ \psi + \varphi' \circ \psi')(x) = 0 \end{cases}$$

A similar calculation as above shows that $\tilde{\rho}$ satisfies $\tilde{\rho} \circ (\varphi \circ \psi + \varphi' \circ \psi') = \rho \circ \varphi \circ \psi + \rho' \circ \varphi' \circ \psi$.

To summarize, we have shown that there exist a FOSD kernel $\tilde{\varphi}$ and a probability kernel $\tilde{\rho}$ that satisfy

$$\tilde{\rho} \circ \tilde{\varphi} \circ (\psi + \psi') = \rho \circ \varphi \circ \psi + \rho' \circ \varphi' \circ \psi' = \phi + \phi' \leq \mu + \mu'.$$

It follows that $\psi + \psi' \preceq_D \mu + \mu'$, as desired. $\square$

The next lemma shows that the random walk kernel also preserves domination.

Lemma 9. *Let σ be a random walk kernel and $\mu, \psi \in \Delta(\Gamma)$ be two positive measures on Γ such that $\psi \preceq_D \mu$. Then $\sigma \circ \psi \preceq_D \sigma \circ \mu$.*

Proof. We first show that if $x, y \in \Gamma$ and $x \leq y$, then

$$\sigma \circ \delta_x \preceq_D \sigma \circ \rho \circ \delta_y. \tag{11}$$

If $x = y$, condition (11) is equivalent to $\sigma \circ \delta_x \preceq_D \sigma \circ \rho \circ \delta_x$. Since ρ is a probability kernel, $\delta_x \preceq_B \rho \circ \delta_x$. Since σ is Blackwell order-preserving, we have $\sigma \circ \delta_x \preceq_B \sigma \circ \rho \circ \delta_x$. Condition (11) follows because domination is implied by Blackwell order dominance.

If $x < y$, define a FOSD kernel $\tilde{\varphi}$ as $\tilde{\varphi}(x') = \delta_y$ for all $x' \leq y$ and $\tilde{\varphi}(x') = \delta_{x'}$ for $x' > y$. Define a probability kernel $\tilde{\rho}$ as $\tilde{\rho} = \sigma \circ \rho$. Since $x < y$ and σ is a random walk kernel, the realizations of $\sigma \circ \delta_x$ lie weakly below y. Therefore,

$\tilde{\varphi} \circ (\sigma \circ \delta_x) = \delta_y$, and thus $\tilde{\rho} \circ \tilde{\varphi} \circ (\sigma \circ \delta_x) = \tilde{\rho} \circ \delta_y = \sigma \circ \rho \circ \delta_y$. Condition (11) follows from the definition of domination.

We are now in position to prove the Lemma. Since $\varphi(x)([x,1]) = 1$, there exists $\beta_y \geq 0$ for all $y \in \Gamma$ with $y \geq x$ such that $\sum_{y \geq x} \beta_y = 1$ and $\varphi(x) = \sum_{y \geq x} \beta_y \delta_y$. Note that $\sigma \circ \rho \circ \varphi \circ \delta_x = \sum_{y \geq x} \beta_y \sigma \circ \rho \circ \delta_y$ and $\sigma \circ \delta_x = \sum_{y \geq x} \beta_y \sigma \circ \delta_x$. It follows from condition (11) and Lemma 8 that

$$\sigma \circ \delta_x \preceq_D \sigma \circ \rho \circ \varphi \circ \delta_x. \tag{12}$$

Since $\psi \preceq_D \mu$, there exist φ and ρ such that $\rho \circ \varphi \circ \psi \leq \mu$. Let $\psi = \sum_x \alpha_x \delta_x$, and hence $\sigma \circ \psi = \sum_x \alpha_x \sigma \circ \delta_x$. Let $\phi \leq \mu$ such that $\rho \circ \varphi \circ \psi = \phi$. Then we can write $\sigma \circ \phi = \sum_x \alpha_x \sigma \circ \rho \circ \varphi \circ \delta_x$. It follows from (12) and Lemma 8 that

$$\sigma \circ \psi \preceq_D \sigma \circ \phi,$$

which, together with the fact that $\sigma \circ \phi \leq \sigma \circ \mu$, implies that $\sigma \circ \psi \preceq_D \sigma \circ \mu$. □

The following lemma shows that domination plays in favor of the sender.

Lemma 10. *Let $\mathbf{X} = (X_t)_{t=1,\ldots}$ and $\mathbf{X}' = (X'_t)_{t=1,\ldots}$ be two random walks on Γ with the same kernel σ. If $\mu'_1 \preceq_D \mu_1$, then the optimal policy under $\mathbf{X}$ yields a higher payoff for the sender than the optimal policy under $\mathbf{X}'$.*

Proof. The proof uses similar considerations as the proof of Theorem 1. Let $\{\nu'_t, \mu'_t\}$ be an optimal interval strategy of the sender for $\mathbf{X}'$. We will show that, if $\mu'_t \preceq_D \mu_t$, then there exists ν_t such that $|\nu_t| = |\nu'_t|$, $\overline{\nu}_t \geq l$ and $\mu'_t - \nu'_t \preceq_D \mu_t - \nu_t$. To see this, note that if $\mu'_t \preceq_D \mu_t$, then there exists φ, ρ, and $\phi \leq \mu_t$ such that $\varphi \circ \rho \circ \mu'_t = \phi \leq \mu_t$. Consider $\nu_t = \varphi \circ \rho \circ \nu'_t$. Since φ is a FOSD kernel and ρ is a probability kernel, we have $\nu_t \leq \phi \leq \mu_t$, $\overline{\nu}_t \geq l$ and $|\nu_t| = |\nu'_t|$. Finally, by construction,

$$\varphi \circ \rho \circ (\mu'_t - \nu'_t) = \varphi \circ \rho \circ \mu'_t - \varphi \circ \rho \circ \nu'_t = \phi - \nu_t \leq \mu_t - \nu_t,$$

which implies $\mu'_t - \nu'_t \preceq_D \mu_t - \nu_t$. By Lemma 9, $\mu_{t+1} = \sigma \circ (\mu_t - \nu_t)$ dominates $\mu'_{t+1} = \sigma \circ (\mu'_t - \nu'_t)$. Therefore, if $\mu'_1 \preceq_D \mu_1$, then for any policy $\{\nu'_t, \mu'_t\}$ under $\mathbf{X}'$, we can construct a policy $\{\mu_t, \nu_t\}_{t=1,\ldots}$ under $\mathbf{X}$ such that $|\nu_t| = |\nu'_t|$ and $\overline{\nu}_t \geq l$. This concludes the proof. □

We call a grid point z_j for $a < j < b$ odd (even) if j is odd (even). For $j \in \{a, b\}$ such that $j \neq \infty, -\infty$, the point z_j will be defined to be both odd and even. We denote by $\Delta^*(\Gamma)$ the set of positive finite measures that are supported on either the even or odd points of the grid. Note that if $\mu \in \Delta^*(\Gamma)$, then $\sigma \circ \mu \in \Delta^*(\Gamma)$.

The following lemma provides an upper bound on how much the sender can improve her persuasion mass by waiting one additional period. For any measure $\mu \in \Delta(\Gamma)$, let ν be the greedy policy with respect to μ, and define $g(\mu) := |\nu|$.

Lemma 11. *For any measure $\mu \in \Delta^*(\Gamma)$, if $\delta \leq \overline{\delta}$ where $\overline{\delta}$ is defined in (8), then*

$$\delta g(\sigma \circ \mu) \leq g(\mu) + \delta g(\sigma \circ (\mu - \nu)).$$

This essentially says that, in a two-period model, if $\delta \leq \overline{\delta}$, then the greedy policy is better than waiting and not revealing any information at the first period and then applying the greedy in the second.

Proof. Since by assumption $l \in (z_0, z_1]$, the lowest point in the support of the greedy measure ν of μ is either z_0 or some $z_j < z_0$. We treat these two cases separately.

Case 1. The lowest point in the support of ν is $z_j < z_0$. In this case, $g(\sigma \circ (\mu - \nu)) = 0$, and thus it is sufficient to show $\overline{\delta} \leq g(\mu)/g(\sigma \circ \mu)$.

Let ν' be the greedy measure for $\sigma \circ \mu$. Then the lowest point in the support of ν' cannot lie strictly below z_{j-1} (above z_{j+1}, respectively), because the conditional expectation of ν' will be strictly below (above, respectively) l. Therefore, the lowest point in the support of ν' is either z_{j-1} or z_{j+1}. We consider these two sub-cases separately.

(i) If the lowest point in the support of ν' is z_{j-1}, then all the mass of ν' that lies strictly above z_{j-1} is contained in $\sigma \circ \nu$. Since $\overline{\sigma \circ \nu} = l$, we must have $\nu' = \sigma \circ \nu$, and hence $|\nu|/|\nu'| = 1$. Therefore, we have $\overline{\delta} \leq 1 = |\nu|/|\nu'| = g(\mu)/g(\sigma \circ \mu)$, as desired.

(ii) If the lowest point in the support of ν' is z_{j+1}, the proof is more involved. Let λ be the submeasure of ν that contains all mass of points lying strictly above z_j. Then we can write $\nu = \alpha \delta_{z_j} + \lambda$, where $\alpha > 0$ is the mass that ν assigns to z_j. It follows from $\overline{\nu} = l$ that

$$\alpha = \frac{\overline{\lambda} - l}{l - z_j} |\lambda|.$$

Since λ contains all mass of points in ν lying strictly above z_j, the lowest point in the support of $\sigma \circ \lambda$ lies (weakly) above z_{j+1} and $\overline{\sigma \circ \lambda} = \overline{\lambda} > l$. Therefore, $\sigma \circ \lambda \leq \nu'$. Furthermore, by assumption, the lowest point in the support of ν' is z_{j+1}, so we can write $\nu' = \beta \delta_{z_{j+1}} + \sigma \circ \lambda$, where $\beta > 0$ is the mass that ν' assigns to z_{j+1}. Again it follows from $\overline{\nu'} = l$ that

$$\beta = \frac{\overline{\sigma \circ \lambda} - l}{l - z_{j+1}} |\sigma \circ \lambda| = \frac{\overline{\lambda} - l}{l - z_{j+1}} |\lambda|.$$

Combining the expressions for α and β, we obtain

$$\frac{g(\mu)}{g(\sigma \circ \mu)} = \frac{|\nu|}{|\nu'|} = \frac{\alpha + |\lambda|}{\beta + |\lambda|} = \frac{l - z_{j+1}}{l - z_j} \frac{\overline{\lambda} - z_j}{\overline{\lambda} - z_{j+1}} \geq \frac{l - z_{j+1}}{l - z_j} \geq \overline{\delta}.$$

Case 2. The lowest point in the support of ν is z_0. For this case, we prove a stronger claim that $g(\mu) + g(\sigma \circ (\mu - \nu)) \geq g(\sigma \circ \mu)$. Again let ν' be the greedy measure for $\sigma \circ \mu$, and let λ be the submeasure that contains all points that lie strictly above z_0 in μ. It must hold that $\sigma \circ \lambda \leq \nu'$. Moreover, since z_0 is the right-most mass of $\sigma \circ (\mu - \lambda)$, ν' must contain all mass that arrives from z_0. Therefore, $\sigma \circ \nu \leq \nu'$. Hence, we have $\sigma \circ \mu \geq \nu' = \sigma \circ \nu + (\nu' - \sigma \circ \nu)$, where $\nu' - \sigma \circ \nu \geq 0$ is positive and $\overline{\nu' - \sigma \circ \nu} = l$. Rearranging yields

$$\sigma \circ (\mu - \nu) = \sigma \circ \mu - \sigma \circ \nu \geq \nu' - \sigma \circ \nu.$$

This implies that $g(\sigma \circ (\mu - \nu)) \geq |\nu' - \sigma \circ \nu| = |\nu'| - |\sigma \circ \nu|$ and therefore

$$g(\mu) + g(\sigma \circ (\mu - \nu)) \geq |\nu| + |\nu'| - |\sigma \circ \nu| = |\nu'| = g(\sigma \circ \mu),$$

where we use the fact that $|\nu| = |\sigma \circ \nu|$. This completes the proof. □

The following corollary essentially shows that the greedy policy is optimal for the case where $T = 2$.

Corollary 2. *For any measure $\mu \in \Delta^*(\Gamma)$, any $\alpha \leq g(\mu)$, and every $\delta \leq \overline{\delta}$, we have*

$$\alpha + \delta g(\sigma \circ (\mu - \nu_\alpha)) \leq g(\mu) + \delta g(\sigma \circ (\mu - \nu)),$$

where $\nu_\alpha \leq \mu$ is the measure associated with the unique interval policy for μ of mass α.

Proof. Let $\mu' = \mu - \nu_\alpha$ and denote ν' to be the greedy measure for μ'. By the interval property of ν_α we have $\nu = \nu_\alpha + \nu'$. Therefore, $g(\mu') = |\nu'| = g(\mu) - \alpha$, and $\mu' - \nu' = \mu - \nu$. Using Lemma 11 for the measure μ', together with the latter two properties we obtain

$$\begin{aligned}\delta g(\sigma \circ \mu') &\leq g(\mu') + \delta g(\sigma \circ (\mu' - \nu')) \\ &= g(\mu) - \alpha + \delta g(\sigma \circ (\mu - \nu)),\end{aligned}$$

thus giving the desired result. □

We are now ready to prove Theorem 2.

Proof of Theorem 2. Fix a discount factor $\delta \leq \overline{\delta}$ and let μ denote the (posterior) distribution of the starting period (i.e., $X_1 \sim \mu$). We prove Theorem 2 first for $T < \infty$ under the assumption that the posterior $\mu \in \Delta^*(\Gamma)$. Denote by $v_t(\mu)$ the value of the sender's problem as a function of the posterior μ and the number of (remaining) periods t. Denote by $\gamma_t(\mu)$ the sender's payoff obtained by following the greedy policy for all t periods, starting with the posterior μ.

We prove by induction on T that $v_T(\mu) = \gamma_T(\mu)$. The proof requires the induction hypothesis to hold for preceding two consecutive periods. Therefore, to start the induction, we need to verify the claim for both $T = 1$ and $T = 2$. But the claim holds for $T = 1$ trivially and for $T = 2$ by Corollary 2.

Now assume that the claim $v_t(\mu) = \gamma_t(\mu)$ holds for $t = T-1, T-2$. We need to show that $v_T(\mu) = \gamma_T(\mu)$. By Theorem 1, we may restrict attention to interval policies. Using the induction hypothesis, we can rewrite the Bellman equation for $v_T(\mu)$ as

$$v_T(\mu) = \max_{\beta \leq g(\mu)} \{\beta + \delta \gamma_{T-1}(\sigma \circ (\mu - \nu_\beta))\}, \tag{13}$$

where ν_β is the measure associated with the unique interval policy for μ of mass $\beta \leq g(\mu)$.

Let $\alpha \leq g(\mu)$ be the maximizer in (13). We write $\mu_2' = \sigma \circ (\mu - \nu_\alpha)$ and let ν_2' be the greedy measure with respect to μ_2'. Let $\mu_2 = \sigma \circ (\mu - \nu)$, where we recall that ν is the greedy measure for μ, and let ν_2 be the greedy measure for μ_2. By Lemma 7, $\mu_2' - \nu_2' \preceq_D \mu_2 - \nu_2$. By Lemma 9, σ preserves the order of domination, so we have $\sigma \circ (\mu_2' - \nu_2') \preceq_D \sigma \circ (\mu_2 - \nu_2)$. By Lemma 10, the domination plays in the favor of the sender, so $v_{T-2}(\sigma \circ (\mu_2' - \nu_2')) \leq v_{T-2}(\sigma \circ (\mu_2 - \nu_2))$, or equivalently by the induction hypothesis, $\gamma_{T-2}(\sigma \circ (\mu_2' - \nu_2')) \leq \gamma_{T-2}(\sigma \circ (\mu_2 - \nu_2))$. This implies that, for any $\delta \leq \overline{\delta}$,

$$\begin{aligned} v_T(\mu) &= \alpha + \delta \gamma_{T-1}(\sigma \circ (\mu - \nu_\alpha)) \\ &= \alpha + \delta g(\mu_2') + \delta^2 \gamma_{T-2}(\sigma \circ (\mu_2' - \nu_2')) \\ &\leq g(\mu) + \delta g(\sigma \circ (\mu - \nu)) + \delta^2 \gamma_{T-2}(\sigma \circ (\mu_2 - \nu_2)) \\ &= \gamma_T(\mu), \end{aligned} \tag{14}$$

where the first equality follows from the induction hypothesis, the inequality uses Corollary 2, and the second and the last equality follow from the definition of the greedy policy. As the reverse inequality $v_T(\mu) \geq \gamma_T(\mu)$ holds by definition, (14) implies that $v_T(\mu) = \gamma_T(\mu)$. This completes the induction step and thus also the proof for the case $T < \infty$.

As for the case $T = \infty$, we let $v_\infty(\mu)$ and $\gamma_\infty(\mu)$ denote the value and the greedy payoff for $T = \infty$. Note that $v_t(\mu) \uparrow v_\infty(\mu)$ and $\gamma_t(\mu) \uparrow \gamma_\infty(\mu)$ as $t \uparrow \infty$. Therefore it must hold that $\gamma_\infty(\mu) = v_\infty(\mu)$ as otherwise we would get that $\gamma_t(\mu) < v_t(\mu)$ for some $t < \infty$, arriving at a contradiction. $\square$

References

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

Robert J Aumann, Michael Maschler, and Richard E Stearns. *Repeated games with incomplete information*. MIT press, 1995.

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

Jacopo Bizzotto, Jesper Rüdiger, and Adrien Vigier. Dynamic persuasion with outside information. *American Economic Journal: Microeconomics*, 13(1):179–194, 2021.

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

Yeon-Koo Che and Johannes Hörner. Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, 133(2):871–925, 2018.

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

Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise. *Journal of Political Economy*, 123(1):215–260, 2015.

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

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

Matteo Escudé and Ludvig Sinander. Slow persuasion. *Theoretical Economics*, 18 (1):129–162, 2023.

Huiyi Guo, Wei He, and Bin Liu. Learning by consuming: Optimal pricing with endogenous information provision. *arXiv preprint arXiv:2209.01453*, 2022a.

Yingni Guo and Eran Shmaya. The interval structure of optimal disclosure. *Econometrica*, 87(2):653–675, 2019.

Yingni Guo, Hao Li, and Xianwen Shi. Optimal discriminatory disclosure. Technical report, Technical Report, Working paper.[1251], 2022b.

Emeric Henry and Marco Ottaviani. Research and the approval process: The organization of persuasion. *American Economic Review*, 109(3):911–955, 2019.

Takakazu Honryo. Dynamic persuasion. *Journal of Economic Theory*, 178:36–58, 2018.

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

Jan Knoepfle and Julia Salmi. Dynamic evidence disclosure: Delay the good to accelerate the bad. *arXiv preprint arXiv:2406.11728*, 2024.

Andrew Kosenko. Algebraic properties of blackwell's order and a cardinal measure of informativeness. *arXiv preprint arXiv:2110.11399*, 2021.

Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the "wisdom of the crowd". *Journal of Political Economy*, 122(5):988–1012, 2014.

Ehud Lehrer and Dimitry Shaiderman. Markovian persuasion. *arXiv preprint arXiv:2111.14365*, 2021.

Laurent Mathevet, David Pearce, and Ennio Stacchetti. Reputation for a degree of honesty. Technical report, Working paper, 2022.

Andrew McClellan. Experimentation and approval mechanisms. *Econometrica*, 90(5):2215–2247, 2022.

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

Robert R Phelps. *Lectures on Choquet's theorem*. Springer, 2001.

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

Moshe Shaked and J George Shanthikumar. *Stochastic orders*. Springer, 2007.

A Ommited Proofs from Sections 2-4

This appendix collects the proofs (for lemmas) omitted in the body of the paper.

Lemma 1. *The sender's optimization problem given in equation (3) has the following equivalent reformulation:*

$$\max_{\{\nu_t\}_{t=1}^T} \sum_{t=1}^{T} |\nu_t| w_t \tag{4}$$

subject to the following recursively defined constraints: $X_1 \sim \mu_1$, $\nu_t \leq \mu_t$ for every $t = 1, \ldots, T$, $\overline{\nu_t} \geq l$ for every $t = 1, \ldots, T$, and $\mu_t := \sigma_{t-1} \circ (\mu_{t-1} - \nu_{t-1})$ for every $t = 2, \ldots, T$.

Proof. We first show that the above sequences of measures are well defined for any given stopping rule. Given a stopping rule τ, we define, for every $t = 1, \ldots, T$, the positive measure $\mu_t = \mathbb{P}_{X_t}(\,\cdot\,|\tau \geq t)\,\mathbb{P}(\tau \geq t)$ corresponding to the unconditional distribution of X_t on the event $\{\tau \geq t\}$. Also, let $\nu_t = \mathbb{P}_{X_t}(\,\cdot\,|\tau = t)\,\mathbb{P}_{X_t}(\tau = t)$ be the measure corresponding to the unconditional distribution of X_t on the event $\{\tau = t\}$. Clearly, $\overline{\nu_t} \geq l$. Moreover, by definition, $\mu_t = \sigma_{t-1} \circ (\mu_{t-1} - \nu_{t-1})$ and $\nu_t \leq \mu_t$ for every $t \geq 1$, and hence $\{\mu_t\}_{t \geq 1}$ and $\{\nu_t\}_{t \geq 1}$ meet the required constraints. Lastly, note that as $\mathbb{P}(\tau = t) = |\nu_t|$ we have $\sum_{t=1}^{T} \mathbb{P}(\tau = t) w_t = \sum_{t=1}^{T} |\nu_t| w_t$, as desired.

Conversely, for every two sequences of positive measures $\{\mu_t\}_{t \geq 1}$, $\{\nu_t\}_{t \geq 1}$ that satisfy the above relations, we will show that there exists a stopping rule τ such that $\mu_t = \mathbb{P}_{X_t}(\,\cdot\,|\tau \geq t)\mathbb{P}(\tau \geq t)$ and $\nu_t = \mathbb{P}_{X_t}(\,\cdot\,|\tau = t)\,\mathbb{P}(\tau = t)$.

To see this, we define the stopping rule recursively such that τ_t depends only on the realization x_t of X_t. Since $\nu_1 \leq \mu_1$, the Radon Nikodym derivative $\frac{d\nu_1}{d\mu_1} : [0,1] \to \mathbb{R}$ satisfies $\frac{d\nu_1}{d\mu_1}(x) \leq 1$ for μ_1 almost every $x \in [0,1]$. By setting $\tau_1 := \frac{d\nu_1}{d\mu_1}$, we have $\nu_1 = \mathbb{P}_{X_1}(\,\cdot\,|\tau = 1)\,\mathbb{P}(\tau = 1)$ as required. We proceed inductively. Assume that we have defined $\{\tau_j\}_{j=1,\ldots,t-1}$ such that $\mu_j = \mathbb{P}_{X_j}(\,\cdot\,|\tau \geq j)\mathbb{P}(\tau \geq j) = \mathbb{P}_{X_j}(\,\cdot\,|\tau > j-1)\mathbb{P}(\tau > j-1)$ and $\nu_j = \mathbb{P}(\tau = j)\mathbb{P}_{X_j}(\,\cdot\,|\tau = j)$ for every $j < t$. Again we can set $\tau_t(x_t) := \frac{d\nu_t}{d\mu_t}$ so that $\nu_t(B) = \int_B \tau_t(x) d\mu_t(x)$ for every Borel set $B \subseteq [0,1]$. We have for any Borel set B,

$$\nu_t(B) = \mathbb{P}(\tau = t, X_t \in B) = \mathbb{P}(\tau = t)\mathbb{P}_{X_t}(X_t \in B|\tau = t). \tag{15}$$

The recursive relations between $\{\mu_t\}_{t \geq 1}$ and $\{\nu_t\}_{t \geq 1}$, coupled with Eq. (15) and the induction assumption yield

$$\begin{aligned}\mu_{t+1} &= \sigma_t \circ (\mu_t - \nu_t) \\ &= \sigma_t \circ (\mu_t - \mathbb{P}_{X_t}(\,\cdot\,|\tau = t)\,\mathbb{P}(\tau = t)) \\ &= \sigma_t \circ (\mathbb{P}_{X_t}(\,\cdot\,|\tau \geq t)\,\mathbb{P}(\tau \geq t) - \mathbb{P}_{X_t}(\,\cdot\,|\tau = t)\,\mathbb{P}(\tau = t)) \\ &= \sigma_t \circ \mathbb{P}_{X_t}(\,\cdot\,|\tau > t)\,\mathbb{P}(\tau > t) \\ &= \mathbb{P}_{X_{t+1}}(\,\cdot\,|\tau \geq t+1)\,\mathbb{P}(\tau \geq t+1).\end{aligned} \tag{16}$$

The combination of Eqs. (15) and (16) completes the induction step. □

Lemma 2. *A probability kernel $\sigma : [0,1] \to \Delta([0,1])$ is Blackwell order preserving if and only if, for every binary supported $\mu \in \Delta([0,1])$,*

$$\sigma \circ \delta_{\overline{\mu}} \preceq_B \sigma \circ \mu.$$

Similarly, a martingale $(X_t)_{t=1,\ldots,T}$ is Blackwell preserving if and only if, for every $t \leq T-1$, the kernel σ_t satisfies the above condition for every binary supported measure μ such that the support of both μ and $\overline{\mu}$ is contained in the support of X_t.

Proof of Lemma 2. The condition in the lemma asserts that, for every $0 \leq y' \leq y \leq y'' \leq 1$, and for every measure $\mu = \alpha\delta_{y'} + (1-\alpha)\delta_{y''}$ with expectation $\overline{\mu} = y$, it holds that $\sigma \circ \delta_y \preceq_B \sigma \circ \mu$. We need to show that this condition implies σ having the Blackwell-order preserving property.

Consider a measure $\nu \preceq_B \mu$. Then there exists a probability kernel $\rho : [0,1] \to \Delta([0,1])$ such that $\rho \circ \nu = \mu$. First, suppose that the kernel $\rho(x)$ has a binary support for any x. That is, $\rho(x) = \alpha_x\delta_{x'} + (1-\alpha_x)\delta_{x''}$ for $0 \leq x' \leq x \leq x'' \leq 1$ such that $\alpha_x x' + (1-\alpha_x)x'' = x$. Thus we can write, for any Borel measurable set $B \subseteq [0,1]$,

$$\mu(B) = \int_{[0,1]} \alpha_x\delta_{x'}(B) + (1-\alpha_x)\delta_{x''}(B)d\nu(x).$$

In particular, $\sigma\circ\nu = \int_{[0,1]} \sigma\circ\delta_x d\nu(x)$ and $\sigma\circ\mu = \int_{[0,1]} \sigma\circ(\alpha_x\delta_{x'}+(1-\alpha_x)\delta_{x''})d\nu(x)$. The result follows since $\sigma \circ \delta_x \preceq_B \sigma \circ (\alpha_x\delta_{x'} + (1-\alpha_x)\delta_{x''})$ for every x.

Next suppose that the kernel ρ has a general support. By definition, $\rho(x) \in \Delta([0,1])$ is a probability distribution on $[0,1]$ with expectation $x \in [0,1]$. The extreme points of all measures on $[0,1]$ with expectation x is the set of all measures with binary support. Therefore, it follows from Choquet's Theorem (see Phelps [2001]) that $\rho(x)$ can be represented as a probability measure over binary support measures for every x. That is, if we let $B_x \subset \Delta([0,1])$ be the set of all binary support measures with expectation x, then there exists $\lambda_x \in \Delta(B_x)$ such that

$$\rho(x) = \int_{B_x} \alpha_x\delta_{x'} + (1-\alpha_x)\delta_{x''}d\lambda_x(x',x'').$$

Therefore, we can write $\mu = \int_{[0,1]} \left(\int_{B_x} \alpha_x \delta_{x'} + (1-\alpha_x)\delta_{x''} d\lambda_x(x', x'', \alpha_x) \right) d\nu(x)$. Since $\sigma \circ \nu = \int_{[0,1]} \sigma \circ \delta_x d\nu(x)$, we can write

$$\sigma \circ \mu = \int_{[0,1]} \sigma \circ \rho(x) = \int_{[0,1]} \left(\int_{B_x} \sigma \circ (\alpha_x \delta_{x'} + (1-\alpha_x)\delta_{x''}) d\lambda_x(x', x'', \alpha_x) \right) d\nu(x).$$

The result follows since for every x and every $\alpha_x \delta_{x'} + (1-\alpha_x)\delta_{x''} \in B_x$ it holds that $\delta_x \preceq_B \sigma \circ (\alpha_x \delta_{x'} + (1-\alpha_x)\delta_{x''})$. □

Lemma 3. *Consider a martingale $\mathbf{X} = (X_t)_{t=1,\ldots,T}$ that is generated by the kernels $(\sigma_t)_{t=1,\ldots,T-1}$ such that for every $t = 1, \ldots, T-1$ the kernel σ_t represent conditionally independent signal, then the martingale is a Blackwell order preserving martingale.*

Proof. We will show that if σ is a probability kernel that represents a conditionally independent signal, then σ is a Blackwell preserving kernel. Consider first the case where $\mu = \delta_x$ for some $x \in (0,1)$ and ν is some mean-preserving spread of μ. In random variable form, there exists a random variable Y such that $\mathbb{E}(Y \,|\, x) = x$, where x is the constant random variable supported on $\{x\}$. By Aumann et al. [1995], there exists a probability kernel $F_\nu : \{0,1\} \to [0,1]$ (referred to as a *state dependent lottery* in Aumann and Maschler) so that $Y = p_x(m) := \mathbb{P}_x(\omega = 1 \,|\, m)$, where $m \in [0,1]$ is the signal whose distribution is generated by x and F_ν. Consider the random variables $X' = \mathbb{P}_x(\omega = 1 \,|\, s)$ and $Y' = \mathbb{P}_x(\omega = 1 \,|\, m, s)$, where s is the signal generated by x and the probability kernel G corresponding to σ. By definition, $X' \sim \sigma \circ \mu$. Also, as the signals m and s may be chosen to be conditionally independent given the state $\omega \in \{0,1\}$, we have for any Borel set $B \subset [0,1]$:

$$\begin{aligned} \mathbb{P}_x[Y' \in B] &= \int_{[0,1]} \mathbb{P}_{p_x(m)} \left[\mathbb{P}_{p_x(m)}[\omega = 1 \,|\, s] \in B \right] \mathrm{d}\nu(p_x(m)) \\ &= \sigma \circ \nu(B), \end{aligned}$$

so that $Y' \sim \sigma \circ \nu$. As the tower property for conditional expectations implies $\mathbb{E}[Y'|X'] = X'$, we deduce that $\sigma \circ \nu$ is a mean-preserving spread of $\sigma \circ \delta_x$, as required.

We move on to the general case. Assume that $\mu \preceq_B \nu$ and let $\tau : [0,1] \to \Delta([0,1])$ be a probability kernel with $\nu = \tau \circ \mu$. Then, for every non-decreasing and concave $f : [0,1] \to \mathbb{R}$ it holds:

$$\begin{aligned}(\sigma \circ \mu)(f) &= \int_{[0,1]} \left(\int_{[0,1]} f(t) \mathrm{d}(\sigma \circ \delta_x)(t) \right) \mathrm{d}\mu(x) \\ &\geq \int_{[0,1]} \left(\int_{[0,1]} f(t) \mathrm{d}(\sigma \circ \tau(x))(t) \right) \mathrm{d}\mu(x) \\ &= (\sigma \circ \nu)(f),\end{aligned} \tag{17}$$

where the first and last equality follow from disintegration formulas, whereas the inequality follows from the fact that $\delta_x \preceq_B \tau(x)$ together with the first part of the proof which implies $\sigma \circ \delta_x \preceq_B \sigma \circ \tau(x)$ for every x. As Eq. (17) is equivalent to $\sigma \circ \mu \preceq_B \sigma \circ \nu$, the proof is complete. □

Lemma 4. *A martingale that is induced by a random walk on a grid is Blackwell order-preserving.*

Proof. Consider the case where $y' = z_j$, $y = z_i$, and $y'' = z_l$ for some $j < i < l$ and $i, j, l \in Z$. As $j \leq i - 1$ and $l \geq i + 1$ we may write:

$$\delta_{z_{i-1}} \preceq_B \frac{z_l - z_{i-1}}{z_l - z_j} \delta_{z_j} + \frac{z_{i-1} - z_j}{z_l - z_j} \delta_{z_l}$$

and

$$\delta_{z_{i+1}} \preceq_B \frac{z_l - z_{i+1}}{z_l - z_j} \delta_{z_j} + \frac{z_{i+1} - z_j}{z_l - z_j} \delta_{z_l}.$$

The above relations, coupled with the definition of σ and simple algebraic manipulations suffice to deduce:

$$\sigma \circ \delta_{z_i} \preceq_B \frac{z_l - z_i}{z_l - z_j} \delta_{z_j} + \frac{z_i - z_j}{z_l - z_j} \delta_{z_l}$$

As $\mu \preceq_B \sigma \circ \mu$ for every probability measure μ, we deduce from the above relation and the fact that $\preceq_B$ is transitive that

$$\sigma \circ \delta_{z_i} \preceq_B \sigma \circ \left\{ \frac{z_l - z_i}{z_l - z_j} \delta_{z_j} + \frac{z_i - z_j}{z_l - z_j} \delta_{z_l} \right\}$$

thus showing that σ preserves Blackwell's order on binary-supported measures, which by Lemma 2 is sufficient to deduce the σ is Blackwell preserving. □

Lemma 5. *Let α be the persuasion probability of the greedy policy with respect to μ_1. For any $0 < \beta \leq \alpha$ there exists an almost surely unique interval policy τ for which $\mathbb{P}(\tau = 1) = \beta$.*

Proof. Let p_0 be the quantile that defines the greedy policy. As mentioned in the text, if we let F be the CDF of μ_1 and U is a uniformly distributed random variable on $[0,1]$, then for $X = F^{-1}(U)$ it holds that $X \sim \mu_1$. It follows that $\mathbb{E}[X|U \geq p_0] = l$. Consider the function $f(p,x) = \mathbb{E}[X|U \in (p,x)]$. Note that over the domain $0 \leq p < x \leq 1$, the function $f(\cdot\,,\cdot)$ is continuous and strictly increasing in both p and x. Let p_1 satisfy $F^{-1}(p_1) = l$. It follows from the properties of f that for every $p_0 < p < p_1$ there exists a unique $x(p)$ such that $f(p, x(p)) = l$. Moreover, $x(p)$ is a continuous function of p. Therefore, the function $x(p) - p$ is continuous, it attained the value α at p_0, and has a left limit 0 at p_1. By the intermediate value theorem, for every $0 < \beta \leq \alpha$, there exists $(p, x(p))$ such that $\mathbb{E}[X|U \in (p, x(p))] = l$ and $x(p) - p = \beta$. Since $x(p) - p$ is strictly decreasing in p, there exists a unique such p for every β. Denote it by p_β. The two quantiles determined by p_β and $x(p_\beta)$ define a unique stopping rule τ (up to measure zero) with $\mathbb{P}(\tau = 1) = \beta$. This concludes the proof. $\square$

Lemma 6. *Let $\nu' \leq \mu_1$ be the measure that corresponds to the interval policy that stops on $[\underline{y}, \overline{y}]$ with $\overline{\nu'} = l$. Let ν'' with $\nu'' \leq \mu_1$ and $\overline{\nu''} = l$ be another stopping rule that satisfies $|\nu''| = |\nu'|$. Then $\nu' \preceq_B \nu''$.*

Proof. Let F' and F'' be the CDFs of $\frac{1}{|\nu'|}\nu'$ and $\frac{1}{|\nu''|}\nu''$, respectively. To show $\nu' \preceq_B \nu''$ it is necessary and sufficient to show that $\int_0^x F'(y)dy \leq \int_0^x F''(y)dy$ for every $x \in [0,1]$. Assume by way of contradiction that $\int_0^x F'(y)dy > \int_0^x F''(y)dy$ for some $x \in [0,1]$. Since $\int_0^{\underline{y}} F'(y)dy = 0$, we have $x > \underline{y}$.

Next, we argue that $x < \overline{y}$. First, we note that $F'(z) = 1$ for every $z \geq \overline{y}$. Thus, assuming to the contrary that $x \geq \overline{y}$, we obtain

$$\begin{aligned}\int_0^1 F'(y)dy &= \int_0^x F'(y)dy + \int_x^1 1dy \\ &> \int_0^x F''(y)dy + \int_x^1 F''(y)dy = \int_0^1 F''(y)dy.\end{aligned}$$

This contradicts the fact that $\int_0^1 F'(y)dy = 1 - \overline{\nu'} = 1 - \overline{\nu''} = \int_0^1 F''(y)dy$, thus establishing $x \in (\underline{y}, \overline{y})$.

We can further assume that $F'(x) \geq F''(x)$ because otherwise we can just decrease the point x until we reach such a point without violating the strict integral inequality. Let F be the CDF of the probability measure μ_1. On the one hand, since $\nu'' \leq \mu_1$ it holds that $F''(x') - F''(x'') \leq F(x') - F(x'')$ for every $x', x'' \in (\underline{y}, \overline{y})$ such that $x' > x''$. On the other hand, by construction $F(x') - F(x'') = F'(x') - F'(x'')$ for every $x', x'' \in (\underline{y}, \overline{y})$ such that $x' > x''$. Therefore, for every $x' \in (\underline{y}, \overline{y})$

such that $x' > x$ it holds that

$$F''(x') = F''(x)+F''(x')-F''(x) \leq F'(x)+F(x')-F(x) = F'(x)+F'(x')-F'(x) = F'(x').$$

Thus $F''(x') \leq F'(x')$ for every $x' \in (\underline{y}, \overline{y})$ such that $x' > x$. This result, together with our assumption of $\int_0^x F'(y)dy > \int_0^x F''(y)dy$, implies that $\int_0^{\overline{y}} F'(y)dy > \int_0^{\overline{y}} F''(y)dy$. Since $F'(x) = 1$ for every $x \geq \overline{y}$, we infer that $\int_0^1 F'(y)dy > \int_0^1 F''(y)dy$, a contradiction. This concludes the proof. □

B Proof of Proposition 1

Proposition 1. *Suppose $T = \infty$ and consider the standard ϵ-grid with $\Gamma = \{n\epsilon : n \in \mathbb{N}_0, n\epsilon < 1\}$. For every $\delta > \sqrt{2}/2$, there exists ϵ' such that, for all ϵ-grid Γ with $l \in \Gamma$ and $\epsilon < \epsilon'$, there is an initial prior $X_0 \in \Delta(\Gamma)$ for which the greedy policy is sub-optimal.*

Proof. Assume by way of contradiction that the greedy policy is optimal for the ϵ-grid (for sufficiently small ϵ, which will be specified later). For every grid point $g \in \Gamma$ and $g < l$ we denote by $v(g)$ the value of the sender (for the entire process) that initiates at g under the greedy policy.

We denote $c = v(l - \epsilon) < 1$. We argue that

$$v(l - 2\epsilon) = c^2 \pm O(\epsilon^2). \tag{18}$$

To see it, we notice that if the random walk that starts at $l - 2\epsilon$ reaches $l - \epsilon$ we can refer to it as if the game terminates and she receives a utility of c. We know that jumping to the next point above $l - \epsilon \to l$ takes an expected discount time c. So, jumping to the next point above $l - 2\epsilon \to l - \epsilon$ should also take an expected discount time of approximately c. Indeed we can couple the two random walks: The one that starts at $l - \epsilon$ and reaches l and the one that starts at $l - 2\epsilon$ and reaches $l - \epsilon$. Failing of this coupling happens with probability $O(\epsilon^2)$; this failure occurs only if the realization of the random walk goes $l - \epsilon \to \epsilon \to l$ without visiting l or 0 in the middle. The coupled process will fail to do the same trajectory because it will move $l - \epsilon \to 0$ and 0 is an observing state.

Using similar arguments one can show that $v(l - 3\epsilon) = c^3 + O(\epsilon^2)$.

By the definition of value, we know that

$$v(l - \epsilon) = \frac{\delta}{2}(1 + v(l - 2\epsilon)) \tag{19}$$

because with probability $\frac{1}{2}$ we will reach l tomorrow and the discounted value would be δ and with probability $\frac{1}{2}$ we will reach $l-2\epsilon$ tomorrow and the discounted value would be $\delta v(l-2\epsilon)$. By combining equations (18) with (19) we get $c = \frac{\delta}{2}(1+c^2) \pm O(\epsilon^2)$ which implies that

$$c = \frac{1-\sqrt{1-\delta^2}}{\delta} \pm O(\epsilon^2). \tag{20}$$

Consider the initial belief X_0 which is distributed as follows: With probability $p = 1 - \frac{\epsilon}{2-2l+\epsilon}$ the belief is $l-2\epsilon$ and with probability $1-p = \frac{\epsilon}{2-2l+\epsilon}$ the belief is 1.

If the sender uses the greedy policy at time $t=1$, she pools together $\frac{p}{4}$ of the mass located at $l-2\epsilon$ with the $1-p$ mass located at 1 and her utility is

$$1 - \frac{3p}{4} + \frac{3p}{4}c^2 \pm O(\epsilon^2),$$

where the term $\frac{3p}{4}c^2$ captures the $\frac{3p}{4}$ mass that remains at $l-2\epsilon$.

If, instead the sender stays mute at time $t=1$ and for time $t \geq 2$ she behaves greedily the following will happen. At time $t=2$ she pools together the $\frac{p}{2}$ mass at $l-\epsilon$ with the $1-p$ mass at 1. This leaves a mass of $\frac{p}{2}$ at $l-3\epsilon$ for future utilization of adoption. In total, her value is

$$\delta(1-\frac{p}{2}) + \delta\frac{p}{2}c^3 \pm O(\epsilon^2).$$

In order for such a deviation from the greedy policy to be profitable we should have

$$\delta(1-\frac{p}{2}) + \delta\frac{p}{2}c^3 > 1 - \frac{3p}{4} + \frac{3p}{4}c^2 \pm O(\epsilon^2)$$

Since $p = 1 - O(\epsilon)$ it is sufficient to have

$$\delta\frac{1}{2} + \delta\frac{1}{2}c^3 > \frac{1}{4} + \frac{3}{4}c^2 \pm O(\epsilon)$$

Using Equation (20) and neglecting the $O(\epsilon^2)$ and $O(\epsilon)$ error terms the inequality above becomes an inequality of δ only. One can verify that for $\delta > \frac{1}{\sqrt{2}}$ we have $\delta\frac{1}{2} + \delta\frac{1}{2}c^3 > \frac{1}{4} + \frac{3}{4}c^2$. Finally, we set ϵ' such that the total sum of all the $O(\epsilon^2)$ and $O(\epsilon)$ error terms along the proof will not exceed the gap $\delta\frac{1}{2} + \delta\frac{1}{2}c^3 - \frac{1}{4} - \frac{3}{4}c^2 > 0$ for every $\epsilon \leq \epsilon'$. In such cases the deviation from the greedy policy is profitable. □

C Existence of a Maximum for Interval Policies

Proposition 2. *Every martingale* $\mathbf{X} = (X_t)_{t=1,\ldots,T}$ *for the sender has an optimal interval policy.*

Proof. We consider the case where $T = \infty$. Led μ be a positive measure on $[0,1]$ with $\mu([0,1]) = r$ and let $D = \{(x,y) \in [0,r]^2 : x \le y\}$. Define a mapping T_μ from D to the set of positive measures over $[0,1]$ by letting $T_\mu(x,y)$ be the measure ν that contains all mass in μ that lies between the quantiles x and y of ν. That is for any measurable function f:

$$\int_{[0,1]} f(z)d\nu(z) = \int_x^y f(F^{-1}(z))dz,$$

where F^{-1} is the inverse of the CDF of μ. Note that, as in the proof of Lemma 5, it follows that for any interval measure $\nu \le \mu$ it holds that $\nu = T_\mu(\underline{q}, \overline{q})$ for some $0 \le \underline{q} \le \overline{q} \le r$.

It readily follows that T_μ is a continuous mapping from D to the set of positive measures that are endowed with the total variation norm. To see this note that if $|x-x'| \le \epsilon$, $|y-y'| \le \epsilon$, and $f : [0,1] \to [-1,1]$, then $\|T_\mu(x,y) - T_\mu(x',y')\|_{TV} \le 2\epsilon$.

As in Lemma 5 we can identify an interval policy $\{(\mu_t, \nu_t)\}_{t=1,\ldots}$ with a sequence of quantiles $\{\underline{q}_t, \overline{q}_t\}_{t=1,\ldots}$ such that the interval measure $\nu_t \le \mu_t$ equals $T_{\mu_t}(\underline{q}_t, \overline{q}_t)$.

Let $\{(\mu_{t,n}, \nu_{t,n})_{t=1,\ldots}\}_{n=1,\ldots,\infty}$ be a sequence of interval policies that attains the supremum in the limit across all interval policies. Let $\{\{\underline{q}_{t,n}, \overline{q}_{t,n}\}_{t=1,\ldots}\}_{n=1,\ldots}$ be the corresponding quantile representation. We assume first that for every t it holds that $\lim_{n\to\infty} \underline{q}_{t,n} = \underline{p}_t$ and $\lim_{n\to\infty} \overline{q}_{t,n} = \overline{q}_t$. Let $\{(\mu_t, \nu_t)\}_{t=1,\ldots}$ be the measure representation of the limit policy. We claim that the policy $\{\underline{p}_t, \overline{p}_t\}_{t=1,\ldots}$ achieves the optimal payoff.

To see this we prove by induction on t that $\lim_{n\to\infty} \mu_{t,n} = \mu_t$, that $\lim_{n\to\infty} \nu_{t,n} = \nu_t$, and that $\overline{\nu_t} = l$.

Note that $\mu_{1,n} = \mu_1$ is the same measure for any n. The fact that $\lim_{n\to\infty} \nu_{1,n} = \nu_1$ follows from the above observation since $\lim_{n\to\infty} \underline{q}_{1,n} = \underline{p}_1$ and $\lim_{n\to\infty} \overline{q}_{1,n} = \overline{q}_1$. Since $\lim_{n\to\infty} \nu_{1,n} = \nu_1$ it must hold that either ν_1 is the zero measure or else $\overline{\nu_1} = l$.

Assume that the claim holds for $t-1$. That is $\lim_{n\to\infty} \mu_{t-1,n} = \mu_{t-1}$, $\lim_{n\to\infty} \nu_{t-1,n} = \nu_{t-1}$, and $\overline{\nu}_{t-1} = l$. Since $\mu_{t,n} = \sigma_{t-1} \circ (\mu_{t-1,n} - \nu_{t-1,n})$ it follows that $\lim_{n\to\infty} \mu_{t,n} = \lim_{n\to\infty} \sigma_{t-1} \circ (\mu_{t-1,n} - \nu_{t-1,n}) = \sigma_{t-1} \circ (\mu_{t-1} - \nu_{t-1}) = \mu_t$. It therefore follows from the fact that $\lim_{n\to\infty} \underline{q}_{t,n} = \underline{p}_t$ and $\lim_{n\to\infty} \overline{q}_{t,n} = \overline{q}_t$ that

$$\lim_{n\to\infty} \nu_{n,t} = \lim_{n\to\infty} T_{\mu_{n,t}}(\underline{q}_{t,n}, \overline{q}_{t,n}) = T_{\mu_t}(\underline{q}_t, \overline{q}_t) = \nu_t.$$

In addition since $\overline{\nu}_{t,n} = l$ we have that either $\nu_t = 0$ or $\overline{\nu}_t = l$ as desired. Since the utility of the sender from each policy is

$$\sum_{t=1}^{\infty} w_t |\nu_{t,n}|$$

we have that the utility converges to the utility of the limit policy $\sum_{t=1}^{\infty} w_t |\nu_t|$. The claim that the limit policy achieves the optimum now readily follows.

We next show that the quantile converges assumption holds without loss. We can take the original sequence of policies $\{(\mu_{t,n}, \nu_{t,n})_{t=1,\ldots}\}_{n=1,\ldots}$ and take a subsequence $\{(\mu_{t,n_{i_1}}, \nu_{t,n_{i_1}})_{t=1,\ldots}\}_{i_1=1,\ldots}$ such that $\lim_{i_1\to\infty} \underline{q}_{1,n_{i_1}} = \underline{p}_1$ and $\lim_{i_1\to\infty} \overline{q}_{1,n_{i_1}} = \overline{q}_1$.

We proceed inductively to get a sequence of refinements such that $\{(\mu_{t,n_{i_k}}, \nu_{t,n_{i_k}})_{t=1,\ldots}\}_{i_k=1,\ldots}$ is a refinement of $\{(\mu_{t,n_{i_{k-1}}}, \nu_{t,n_{i_{k-1}}})_{t=1,\ldots}\}_{i_2=1,\ldots}$ and $\lim_{i_2\to\infty} \underline{q}_{k,n_{i_k}} = \underline{p}_k$ and $\lim_{i_k\to\infty} \overline{q}_{k,n_{i_2}} = \overline{q}_k$. It is now easy to see that the diagonal subsequence has the desired properties. □